



Maturation de sites métalliques de protéines par les protéines à radical S-Adénosyl-L-méthionine et la machinerie de fabrication des centres fer-soufre

Elodie Marinoni

► To cite this version:

Elodie Marinoni. Maturation de sites métalliques de protéines par les protéines à radical S-Adénosyl-L-méthionine et la machinerie de fabrication des centres fer-soufre. Autre [q-bio.OT]. Université de Grenoble, 2011. Français. NNT : 2011GRENV064 . tel-00680023

HAL Id: tel-00680023

<https://theses.hal.science/tel-00680023>

Submitted on 17 Mar 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ DE GRENOBLE

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Biologie structurale et Nanobiologie**

Arrêté ministériel : 7 août 2006

Présentée par

Elodie N. MARINONI

Thèse dirigée par **Juan Carlos FONTECILLA CAMPS** et
co-dirigée par **Yvain NICOLET**

Thèse préparée au sein de l'**Institut de Biologie Structurale**
Jean-Pierre Ebel (IBS) CEA/CNRS/UJF, Groupe
Métalloprotéines
dans l'**École Doctorale Chimie et Sciences du Vivant**

**Maturation de sites métalliques
de protéines par les protéines à
radical S-Adénosyl-L-
méthionine et la machinerie de
fabrication des centres fer-
soufre**

Thèse soutenue publiquement le **9 décembre 2011**,
devant le jury composé de :

Mme Hélène PUCCIO

Directrice de recherche INSERM, Rapporteur

Mme Sandrine OLLAGNIER DE CHOUDENS

Chargée de recherche CNRS, Examinatrice

M. Juan Carlos FONTECILLA CAMPS

Directeur de recherche CEA, Directeur de thèse

M. Yvain NICOLET

Chargé de recherche CEA, Co-directeur de thèse

M. Luis RUBIO

Chargé de recherche Université de Madrid, Rapporteur

M. David PIGNOL

Directeur de recherche CEA, Président du jury



Remerciements

Je souhaite tout d'abord remercier Juan Fontecilla, mon directeur de thèse, de m'avoir accueillie dans son laboratoire durant mes années de thèse. Merci à lui pour sa patience, son envie de me former à la recherche et ses discussions scientifiques. J'ai beaucoup apprécié son envie de partager sa très grande passion pour la science. Il a toujours été très à cœur sur les projets sur lesquels j'ai travaillé.

Je remercie Yvain Nicolet, mon co-directeur de thèse, pour son aide, la formation qu'il m'a donnée. Il m'a expliqué beaucoup de choses, qu'elles soient techniques ou théoriques. Nous avons aussi eu beaucoup d'espoir en beaucoup de choses, par exemple en allant au synchrotron ensemble en espérant obtenir un super jeu de données, et beaucoup de défaites, je le remercie d'avoir été là dans ces moments-là.

Je remercie Hélène Puccio et Luis Rubio d'avoir accepté d'être mes rapporteurs, Sandrine Ollagnier-de-Choudens et David Pignol de faire aussi partie de mon jury.

Je remercie toutes les personnes qui ont travaillé sur mes projets.

Je remercie Patricia Amara d'avoir toujours été là pour m'aider, pour sa bonne humeur, nos discussions, ses encouragements.

Je remercie Lydie Martin qui m'a vraiment aidée durant les mois de mon arrivée au laboratoire, elle a été très patiente et m'a appris beaucoup de choses. Merci à elle pour sa bonne humeur, nos discussions.

Je remercie Marina Iannello, ma fidèle collègue de bureau, qui a pu assister à tous les grands moments de ma thèse. Merci à elle pour son soutien, sa bonne humeur et nos discussions.

Je remercie plus généralement tous les membres du laboratoire, pour leur accueil, leur bonne humeur et par conséquent la très bonne ambiance qui y règne. J'ai toujours été très heureuse de venir y travailler. Je m'excuse de ne pas avoir nommé tout le monde, mais sinon je crois que ma thèse n'aurait été constituée que de remerciements.

Je remercie les membres de ma famille qui m'ont toujours soutenue durant ma thèse. Merci à mes parents de m'avoir toujours encouragée à continuer mes études. Merci à Florian d'avoir toujours été là pour me remonter le moral et m'avoir encouragée, malgré mes grosses déceptions quand je n'avais toujours pas de résultats, ou la dure période de rédaction.

Sommaire
et
Liste des figures
et tableaux

Sommaire

ABREVIATIONS	13
INTRODUCTION	17
I) Les centres FeS.....	19
I.1) Nature et fonction des centres FeS	19
I.1.1) Différentes formes de centres FeS.....	19
I.1.2) Quelques fonctions des centres FeS.....	21
I.1.2.1) Le transfert d'électrons.....	21
I.1.2.2) La liaison à un substrat et son activation.....	22
I.1.2.3) La détection de conditions environnementales et régulation de l'expression de gènes.....	24
I.2) Les hydrogénases	25
II) Fabrication des centres FeS	26
II.1) Maturation de la nitrogénase et découverte des machineries ISC et SUF.....	27
II.2) La machinerie ISC	29
II.2.1) Exemples d'opérons codant pour les machineries bactériennes de fabrication des centres FeS	29
II.2.2) Les protéines IscS et IscU.....	30
II.2.2.1) La protéine IscS	30
II.2.2.2) La protéine IscU.....	33
II.2.2.3) Assemblage de centres [2Fe-2S] ou [4Fe-4S] sur le domaine N-terminal de NifU et IscU	35
II.3) Assemblage des centres FeS chez les eucaryotes	36
II.4) Identification d'une machinerie de maturation de l'hydrogénase à [Fe].....	38
III) Les protéines à radical SAM	42
III.1) Le cœur ($\beta\alpha$) ₆ des PRS.....	42
III.2) Des structures spéciales pour des PRS	44
III.3) Le centre [4Fe-4S] et la liaison de la SAM.....	45
III.4) Fonction et réactivité des PRS	48
III.5) Sous-famille de PRS ($\beta\alpha$) ₈	49
MATERIEL ET METHODES.....	55
I) Biologie moléculaire.....	57
I.1) Clonage dans le plasmide pEtDuet-1 : plasmides pETMK0016Duet-1 et pETMJ0785Duet-1	57
I.1.1) Amplification des gènes d'intérêt	57
I.1.2) Digestion enzymatique de l'insert et du plasmide	59
I.1.3) Traitement à la phosphatase.....	59
I.1.4) Ligation de l'insert et du plasmide digérés.....	59
I.1.5) Transformation du produit de ligation.....	59
I.1.6) Test PCR sur colonies et extraction de plasmide	60
I.2) Mutant C-terminal.....	60
I.3) Co-transformation de pETMK0016Duet-1, pETMK0016_cut_Duet-1, pETMJ0785Duet-1 avec pRSFIscDuet-1 ou pRSFIscMetKDuet-1	61
II) Biochimie	63
II.1) Protéine HmdB.....	63
II.1.1) Surexpression	63
II.1.2) Purification de la protéine HmdB de Methanopyrus kandleri.....	64
II.1.3) Essai de purification de la protéine HmdB de Methanococcus jannaschii	65
II.2) Complexe IscSU.....	66

II.2.1)	Production et purification de la protéine AflscS et du complexe AflscSU	66
II.2.2)	Calculs théoriques.....	66
II.3)	Dosage au rose bengal et en fer	67
II.3.1)	Dosage au rose bengal.....	67
II.3.2)	Dosage en fer	68
II.4)	Analyse par électrophorèse sur gel de polyacrylamide	69
III)	Cristallogénèse et cristallographie des protéines.....	69
III.1)	Cristallogénèse	69
III.2)	Cristallographie aux rayons X.....	72
III.2.1)	La diffraction des rayons X par un cristal.....	73
III.2.2)	L'enregistrement et le traitement des données de diffraction.....	76
III.2.3)	Qualité d'un jeu de données	77
III.2.4)	Estimation du nombre de molécules par unité asymétrique.....	78
III.2.5)	Remplacement moléculaire.....	78
III.2.6)	La diffusion anormale	81
III.2.7)	La résolution de la structure de la protéine	82
III.2.7.1)	La modification de densité	82
III.2.7.2)	La construction et l'affinement de la structure	84
III.2.7.3)	Les critères de suivi de l'affinement	85
III.2.7.4)	Les cartes de densité électronique	86
IV)	Spectroscopies	87
IV.1)	Séquençage N-terminal de protéine.....	87
IV.2)	Spectroscopie UV-visible.....	87
IV.3)	Spectroscopie de résonance paramagnétique électronique (RPE)	88
IV.4)	Spectroscopie Raman de résonance.....	89
	RESULTATS ET DISCUSSION.....	91
I)	Etude structurale de la protéine HmdB	93
I.1)	Clonage du gène, expression et purification de HmdB native de <i>Methanopyrus kandleri</i> (MkHmdB)	94
I.2)	Essais de cristallisation.....	100
I.3)	Résolution de structure de HmdB.....	102
I.4)	Analyse structurale de la protéine HmdB.....	105
I.4.1)	Résidus non visibles dans sa structure.....	105
I.4.1)	Les limites du modèle de la protéine HmdB	108
I.4.2)	Agencement spatial des motifs conservés des protéines à radical SAM dans MkHmdB	110
I.5)	Empilement cristallin de MkHmdB.....	116
I.6)	Elaboration d'un mutant C-terminal et changement d'organisme	118
I.7)	Conclusion et perspectives.....	122
II)	Etude structurale du complexe IscS-IscU	124
II.1)	Gènes d' <i>Archaeoglobus fulgidus</i> codant pour IscS et IscU	125
II.2)	Structure de AflscS	126
II.2.1)	Caractérisation de AflscS	126
II.2.2)	Cristallisation de AflscS	127
II.2.3)	Structure de AflscS.....	128
II.2.3.1)	Boucle contenant la cystéine conservée	129
II.2.3.2)	Site de fixation du PLP.....	132
II.3)	Structure du complexe AflscSU ₂	133
II.3.1)	Caractérisation de AflscSU	133
II.3.2)	Cristallisation du complexe AflscSU.....	137
II.3.3)	Résolution de structure de AflscSU	139
II.3.4)	Structure de IscSU exposée à l'air (AE).....	140
II.3.5)	Structure de IscSU traitée au DTT.....	144
II.3.6)	Interface entre IscS et IscU	145

II.3.7)	Forme UT	148
II.3.8)	Comparaison des formes DT/AE/UT d'AflscSU	149
II.3.9)	Calculs théoriques.....	150
II.3.10)	Spectres RPE et Raman	153
II.3.11)	Proposition d'un mécanisme de formation des centres FeS.....	155
II.4)	Conclusion et perspectives	158
BIBLIOGRAPHIE		161

Liste des figures

INTRODUCTION

Figure I-1 La diversité des centres FeS.....	20
Figure I-2 Exemple de coordination non commune d'un centre [2Fe-2S] : cas de HydE.....	21
Figure I-3 Réaction catalysée par l'aconitase.....	22
Figure I-4 Ligands citrate ou isocitrate liés à la m-aconitase.....	23
Figure I-5 Liaison à un substrat et son activation.....	24
Figure I-6 Sites actifs des 3 types d'hydrogénase.....	26
Figure II-1 Réaction catalysée par les cystéines désulfurases.....	28
Figure II-2 Organisation des gènes dans des clusters de gènes <i>nif</i> , <i>isc</i> , <i>suf</i> et <i>csd</i>	30
Figure II-3 Alignement multiple de séquences de cystéines désulfurases du groupe I.....	32
Figure II-4 Superposition de structures de cystéines désulfurases.....	32
Figure II-5 Alignement multiple de séquences de protéines IscU et d'homologues.....	34
Figure II-6 Structure d'EclscU.....	35
Figure II-7 Domaines de NifU.....	35
Figure II-8 Réaction catalysée par Hmd.....	38
Figure II-9 Cofacteur FeGP d'Hmd.....	39
Figure II-10 Site actif d'Hmd.....	39
Figure II-11 Gènes adjacents au gène codant pour la protéine Hmd.....	41
Figure III-1 Réaction d'initiation radicalaire des protéines à radical SAM.....	42
Figure III-2 Structures de protéines à radical SAM.....	44
Figure III-3 Superposition de structures connues des PRS.....	44
Figure III-4 Structure des homodimères Dph2 et ThiC.....	45
Figure III-5 Motif GGE.....	46
Figure III-6 Motif GxIxGxxE.....	46
Figure III-7 Alignement structural de PRS.....	47
Figure III-8 Position du substrat et de la SAM dans les PRS.....	48
Figure III-9 Structures des PRS ($\beta\alpha$) ₈ BioB et HydE.....	49
Figure III-10 Alignement structural des protéines HydE et BioB.....	50
Figure III-11 Réactions catalysées par les protéines BioB et PylB.....	50
Figure III-12 Réactions catalysées par les protéines ThiH, HydG, NosL et la FO synthase.....	52

MATERIEL ET METHODES

Figure I-1 Insertion des gènes MK0016 et MJ0785 dans le plasmide pETDuet-1.....	57
Figure I-2 Représentation schématique des plasmides pRSFIscDuet-1 et pRSFIscMetKDuet-1.....	62
Figure III-1 Technique de la diffusion de vapeur.....	70
Figure III-2 Diagramme de phase et chemins expérimentaux pour la technique de diffusion de vapeur....	72
Figure III-3 Phénomène de diffraction selon les lois de Laue (a) et de Bragg (b).....	75
Figure III-4 Loi de Friedel faussée en présence de diffusion anormale.....	82
Figure III-5 Procédure de modification de densité électronique.....	83

RESULTATS ET DISCUSSION

Figure I-1 Réaction catalysée par Hmd.....	93
Figure I-2 Site actif d'Hmd.....	93
Figure I-3 Protéolyse de la protéine <i>MkHmdB</i>	95
Figure I-4 Alignement de séquences de la protéine HmdB de <i>Methanococcus maripaludis</i> S2 et <i>Methanopyrus kandleri</i>	96
Figure I-5 Expression de la protéine <i>MkHmdB</i> native.....	97
Figure I-6 Echantillon protéique final.....	98
Figure I-7 HmdB avant et après reconstitution.....	99
Figure I-8 Cristaux de HmdB.....	101

Figure I-9 Résolution de la structure de <i>MkHmdB</i> .	105
Figure I-10 Structures des protéines à radical SAM ($\beta\alpha$) ₈ .	106
Figure I-11 Alignement structural des protéines <i>MkHmdB</i> , <i>EcBioB</i> , <i>TmHydE</i> .	107
Figure I-12 Limites du modèle de la protéine HmdB.	108
Figure I-13 Motif de fixation du centre [4Fe-4S].	110
Figure I-14 Motif « GxIxGxxE ».	111
Figure I-15 Motif « GGE ».	112
Figure I-16 Motif « CCNLET(ou S)/R » de HmdB, YNHNLDT/R de BioB et LLRHET/R de HydE.	113
Figure I-17 Alignement multiple de séquences de protéines HmdB.	115
Figure I-18 Empilement cristallin de <i>MkHmdB</i> .	116
Figure I-19 Empilement cristallin de <i>MkHmdB</i> , boucles non visibles et superposition sur les structures de BioB et HydE.	118
Figure I-20 Surface de <i>MkHmdB</i> .	118
Figure I-21 Expression de HmdB mutante.	119
Figure I-22 HmdB après purification.	120
Figure I-23 Reconstitution de HmdB mutée.	120
Figure I-24 Expression de HmdB de <i>Methanococcus jannaschii</i> .	121
Figure II-1 Gènes d' <i>Archaeoglobus fulgidus</i> codant pour IscS et IscU.	125
Figure II-2 Spectre UV-visible de <i>AflscS</i> .	126
Figure II-3 Cristaux de <i>AflscS</i> .	127
Figure II-4 Structure de <i>AflscS</i> .	129
Figure II-5 Homologues d' <i>AflscS</i> .	130
Figure II-6 Alignement de séquence multiple avec Clustal W2 (Larkin, Blackshields et al. 2007) de plusieurs IscS et homologues structuraux (Dali-server).	131
Figure II-7 Site de fixation du PLP dans <i>AflscS</i> .	132
Figure II-8 Mécanisme réactionnel des cystéines désulfurase.	133
Figure II-9 Chromatogrammes d' <i>AflscSU</i> injectée sur Superdex 200 HR 10/30.	135
Figure II-10 Gel natif du complexe <i>AflscSU</i> .	135
Figure II-11 SDS-PAGE (non réducteur) du complexe <i>AflscSU</i> .	136
Figure II-12 Formule chimique développée du dithiothreitol.	136
Figure II-13 Spectres UV-visible d' <i>AflscSU</i> .	137
Figure II-14 Cristaux du complexe <i>AflscSU</i> .	139
Figure II-15 Hétérotétramère du complexe AE- <i>AflscSU</i> ₂ .	141
Figure II-16 Site d'assemblage du centre FeS dans IscU dans le complexe AE <i>AflscSU</i> .	142
Figure II-17 Complexe IscSU exposé à l'air (AE).	143
Figure II-18 Hétérotétramère de DT- <i>AflscSU</i> ₂ .	144
Figure II-19 Complexe IscSU traité au DTT (DT).	145
Figure II-20 Comparaison des formes AE/DT d' <i>AflscSU</i> et <i>EclscSU</i> .	146
Figure II-21 Interface des complexes DT <i>AflscSU</i> et <i>EclscSU</i> .	147
Figure II-22 Site de fabrication des centres FeS d'IscU de la forme UT.	149
Figure II-23 Superposition de la forme DT sur le modèle MQ/MM optimisé où l'Ala35 a été substituée par l'Asp.	151
Figure II-24 Modèles MQ/MM.	152
Figure II-25 Interface d'IscS/IscU de la forme DT.	152
Figure II-26 « Pince » Cys33-Asp35.	153
Figure II-27 Spectres Raman.	154
Figure II-28 Différentes conformations de la boucle contenant la cystéine active.	155
Figure II-29 Représentation de la surface du complexe IscSU.	156

Liste des tableaux

INTRODUCTION

Tableau II-1 Composants impliqués dans la biogenèse des centres FeS.....	37
Tableau II-2 Maladies associées à aux protéines à centres FeS et leur biogenèse.....	37

MATERIEL ET METHODES

Tableau I-1 Amorces utilisées pour insérer HmdB dans le pETDuet-1.....	58
Tableau I-2 Amplification des gènes par PCR. Mélange réactionnel et programme.....	58
Tableau I-3 Amorces utilisées pour muter HmdB au codon 336.....	60

RESULTATS ET DISCUSSION

Tableau I-1 Statistiques de collecte, traitement de données et d'affinement pour <i>MkHmdB</i>	102
Tableau I-2 Superposition de la protéine HmdB sur BioB ou HydE.....	108
Tableau I-3 Comparaison de la conformation de résidus situés sur une partie intermédiaire entre le haut et le bas du tonneau des structures de HmdB, BioB et HydE.....	109
Tableau II-1 Statistiques de collecte, traitement de données et d'affinement pour <i>AflscS</i>	128
Tableau II-2 Statistiques de collection, traitement de données et d'affinement pour le complexe <i>AflscSU</i>	140
Tableau II-3 Bandes Raman d'un centre [2Fe-2S].....	155

Abréviations

T _m	température de fusion
LB	Luria Broth
P _b	paire de bases
Kan	Kanamycine
Amp	Ampicilline
TB	Terrific Broth
DO	densité optique
IPTG	isopropyl β-D-1-thiogalactopyranoside
EDTA	acide éthylène diamine tétracétique
DTT	dithiothréitol
AdoMet ou SAM	<i>S</i> -Adénosyl-L-méthionine
SAH	<i>S</i> -Adénosyl-L-homocystéine
BAG	boite à gants en condition anaérobie
CV	volume de colonne
U	unité
MPD	2-méthyl-2,4-pentanediol
PCR	Polymerase Chain Reaction
PEG	polyéthylène glycol
PLP	phosphate de pyridoxal
PRS	protéine à radical SAM
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
ATP	adénosine triphosphate
ESRF	European Synchrotron Radiation Facility (Grenoble)
pI	point isoélectrique

Introduction

I) Les centres FeS

I.1) Nature et fonction des centres FeS

Les centres fer-soufre ou FeS sont des cofacteurs protéiques inorganiques composés entre autres d'atomes de $\text{Fe}^{2+/3+}$ et S^{2-} . Ils sont très utilisés, se trouvent dans les 3 règnes du vivant et ont des rôles divers (Johnson, Dean et al. 2005) :

- ✓ de transfert d'électrons,
- ✓ de liaison à un substrat et son activation,
- ✓ de régulation d'expression de gènes,
- ✓ de donneur de soufre,
- ✓ structural.

Dans les 110 protéines contenant un centre FeS déjà isolées chez *E.coli.*, 50% effectuent une fonction de transfert d'électrons (Fontecave 2006).

I.1.1) Différentes formes de centres FeS

Les centres FeS les plus communs sont les centres [2Fe-2S] rhombique, [4Fe-4S] cubique et [3Fe-4S] (Figure I-1 (a), (b) et (c)). Le centre FeS le plus simple est l'atome de Fe ([1Fe-0S]) de la rubrédoxine, coordonné par 4 cystéines de la protéine.

Des centres FeS plus complexes ont été identifiés, dont voici les caractéristiques:

- ✓ ceux au nombre d'atomes de fer et soufre plus important que dans les centres communs, tel que l'agrégat [8Fe-7S] de la nitrogénase, enzyme qui fixe l'azote dans les organismes azotrophes, c'est-à-dire catalyse la formation d'ammoniaque à partir d'une molécule d'azote. Il est appelé « P-cluster » (Figure I-1 (f)) et est en fait composé de 2 centres [4Fe-3S] cubiques connectés par un atome de soufre central (code PDB 1m1n, (Einsle, Tezcan et al. 2002));
- ✓ des complexes composés de Fe, S et d'autres métaux tels que le Ni ou le Mo. Le « C-cluster » de la CO-déshydrogénase Ni dépendante (Figure I-1 (d)) (code PDB 1oao, (Darnault, Volbeda et al. 2003)) contient une unité [3Fe-4S] liée à un atome de fer et un atome de nickel. Le « FeMo-co » de la nitrogénase (Figure I-1 (g)) est composé d'un atome de Mo, un ion N, 7Fe et 9S (code PDB 1m1n, (Einsle, Tezcan et al. 2002)) ;

- ✓ des centres $[4\text{Fe-4S}]$ liant un autre centre métallique par un de leurs atomes de fer.
- Dans le « A-cluster » de l'acétyl-coenzymeA synthétase (Figure I-1 (e)) (code PDB 1aoa, (Darnault, Volbeda et al. 2003)), un atome de fer du centre $[4\text{Fe-4S}]$ lie un centre à 2 atomes de Ni. Dans la sulfite réductase un atome de fer du centre $[4\text{Fe-4S}]$ lie un sirohème (Figure I-1 (h)) ((code PDB 3mm5, (Parey, Warkentin et al. 2010)). Dans les hydrogénases à $[\text{Fe-Fe}]$ le centre $[4\text{Fe-4S}]$ lie un centre à 2 atomes de Fe : ce site sera décrit plus en détails plus tard.

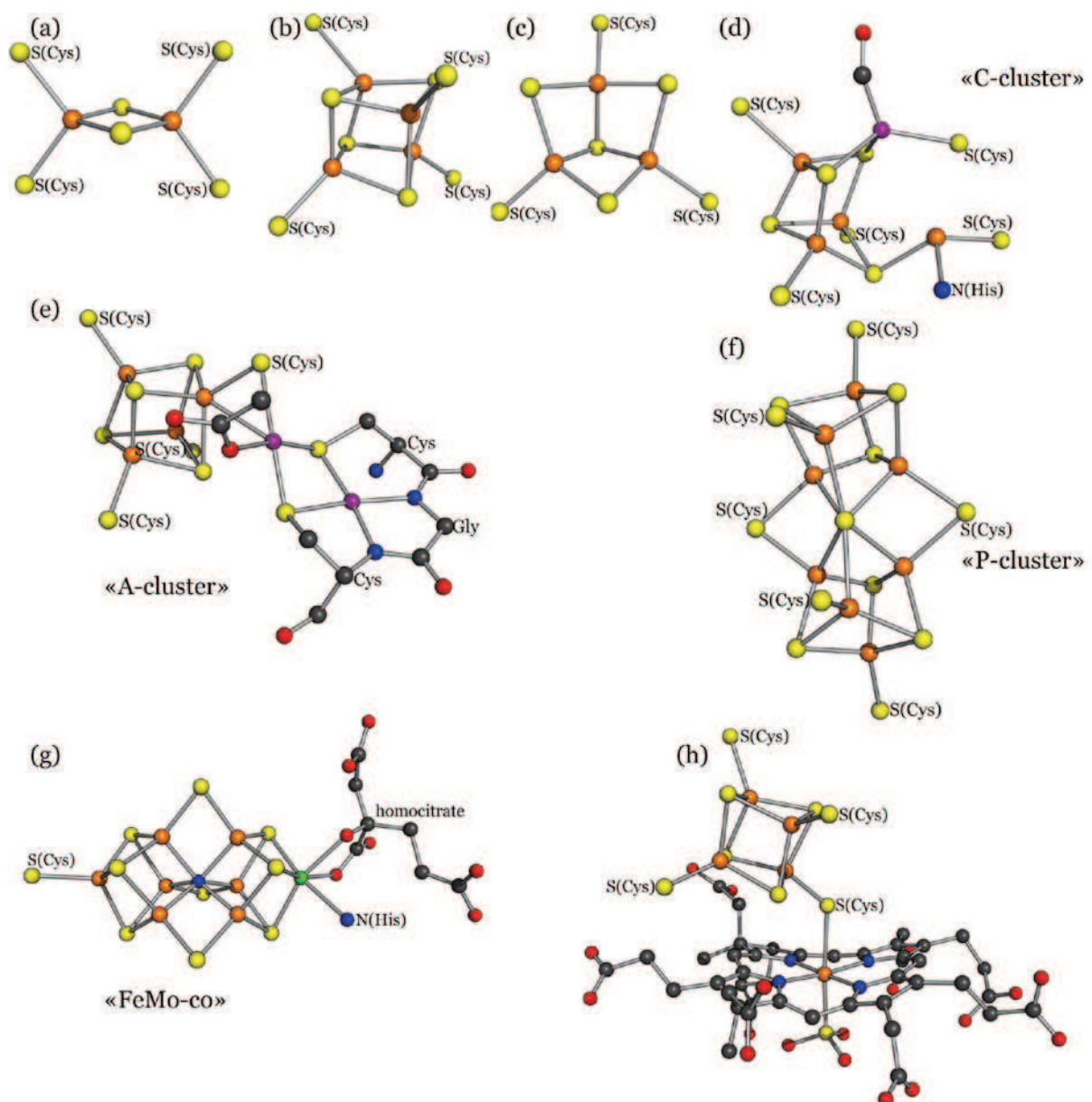


Figure I-1 La diversité des centres FeS.

En jaune : atomes de soufre, orange : fer, bleu : azote, noir : carbone, vert : molybdène, rouge : oxygène, violet : nickel.

(a) centre [2Fe-2S], (b) centre [4Fe-4S], (c) centre [3Fe-4S], (d) « C-cluster » de la CO-déhydrogénase dépendante au Ni (code PDB 1oao, (Darnault, Volbeda et al. 2003)), (e) « A-cluster » de l'acétyl-coenzymeA synthase (code PDB 1oao, (Darnault, Volbeda et al. 2003)), (f) « P-cluster » de la nitrogénase (code PDB 1m1n, (Einsle, Tezcan et al. 2002)), (g) « FeMo-co » de la nitrogénase (code PDB 1m1n, (Einsle, Tezcan et al. 2002)) et (h) centre de la sulfite réductase (code PDB 3mm5, (Parey, Warkentin et al. 2010)).

L'intégration d'un centre FeS à une protéine se fait classiquement par une coordination des atomes de Fe aux résidus Cys ou His de la protéine mais aussi aux ligands Asp, Arg, Ser, Gln et CO ou CN^- . Par exemple la protéine HydE possède un centre [2Fe-2S] dont la coordination varie (Figure I-2). Il a été observé qu'en solution le centre est censé être un [4Fe-4S] (Rubach, Brazzolotto et al. 2005). Il y a probablement un problème d'empilement cristallin qui empêche la cristallisation avec un centre [4Fe-4S]. Dans les cristaux, ses ligands sont trois cystéines et une arginine, qui peuvent faire des relais avec une molécule d'eau (Figure I-2 (a)) ou des atomes de soufre (Figure I-2 (b)) (PDBs 3cix et 3iiz ; (Nicolet, Rubach et al. 2008) et (Nicolet, Amara et al. 2009)).

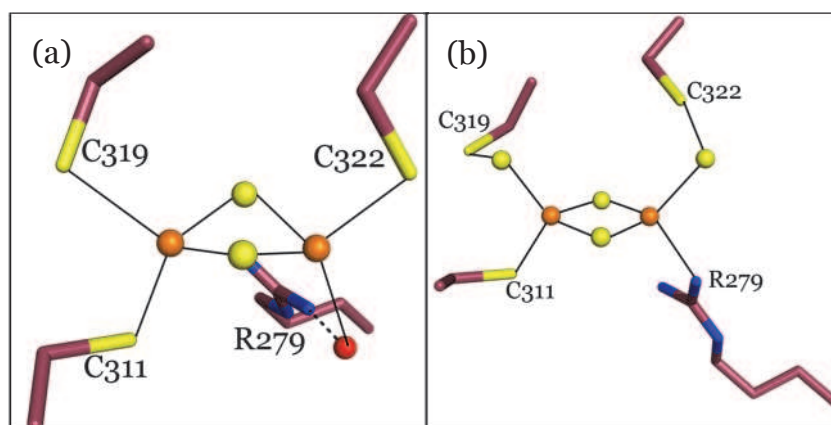


Figure I-2 Exemple de coordination non commune d'un centre [2Fe-2S] : cas de HydE.

En orange, atomes de fer ; en rouge, molécule d'eau. (a) Code PDB 3cix (Nicolet, Rubach et al. 2008). (b) Code PDB 3iiz (Nicolet, Amara et al. 2009).

I.1.2) Quelques fonctions des centres FeS

I.1.2.1) Le transfert d'électrons

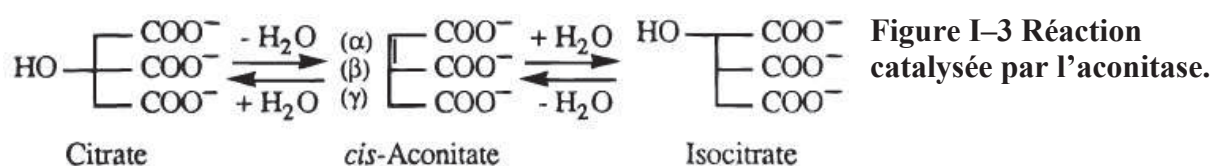
La première fonction découverte des centres FeS est celle de transfert d'électrons. C'est aussi la fonction la plus exercée par ces centres. En effet, les centres sont composés d'atomes de $\text{Fe}^{2+/3+}$ et S^{2-} , le potentiel redox du couple $\text{Fe}^{2+}/\text{Fe}^{3+}$ étant dans le domaine des potentiels

redox d'une cellule. Le centre FeS a un potentiel redox qui varie selon son environnement protéique (Meyer 2008). Par exemple la rubredoxine et les ferrédoxines sont des transporteurs d'électrons. On trouve aussi des chaînes de transfert d'électrons entre différents centres FeS dans les complexes de la chaîne respiratoire. Le "P-cluster" de la nitrogénase a cette fonction.

I.1.2.2) La liaison à un substrat et son activation

La disponibilité d'un site dans un centre FeS pour la liaison à un substrat et son activation peut se faire de différentes manières.

La première façon est la fixation d'un centre [4Fe-4S] par trois cystéines. Avec 3 cystéines, un des atomes de Fe n'aura que 3 ligands S^{2-} et un site vacant. Ce site sera occupé par un ligand, par exemple une molécule d'eau, qui sera échangeable. Par exemple, la fonction primaire de l'aconitase est de catalyser, par son centre [4Fe-4S], la réaction non redox stéréospécifique de déshydratation/réhydratation du citrate en isocitrate, via le *cis*-aconitate (Figure I-3). Ces deux réactions sont réversibles mais la formation d'isocitrate est favorisée. Elles font partie des étapes du cycle de l'acide tricarboxylique (cycle de Krebs). A l'équilibre, le citrate est présent à 88%, l'isocitrate à 8% et le *cis*-aconitate à 4%. L'aconitase a été découverte dans les années 1930 et a été ensuite très étudiée. La revue de (Beinert, Kennedy et al. 1996) donne une vue très large sur les travaux qui ont été faits sur cette protéine.



Comme mentionné plus haut, l'aconitase possède un centre [4Fe-4S] coordonné par 3 cystéines. Un atome de fer est libre et coordine le substrat ainsi qu'une molécule d'eau nécessaire à la catalyse. Le citrate est lié à ce fer libre par le carboxylate et groupement hydroxyle en C β (Figure I-4 (a)). L'isocitrate est coordonné de la même manière au centre [4Fe-4S] mais en position C α (Figure I-4 (b)).

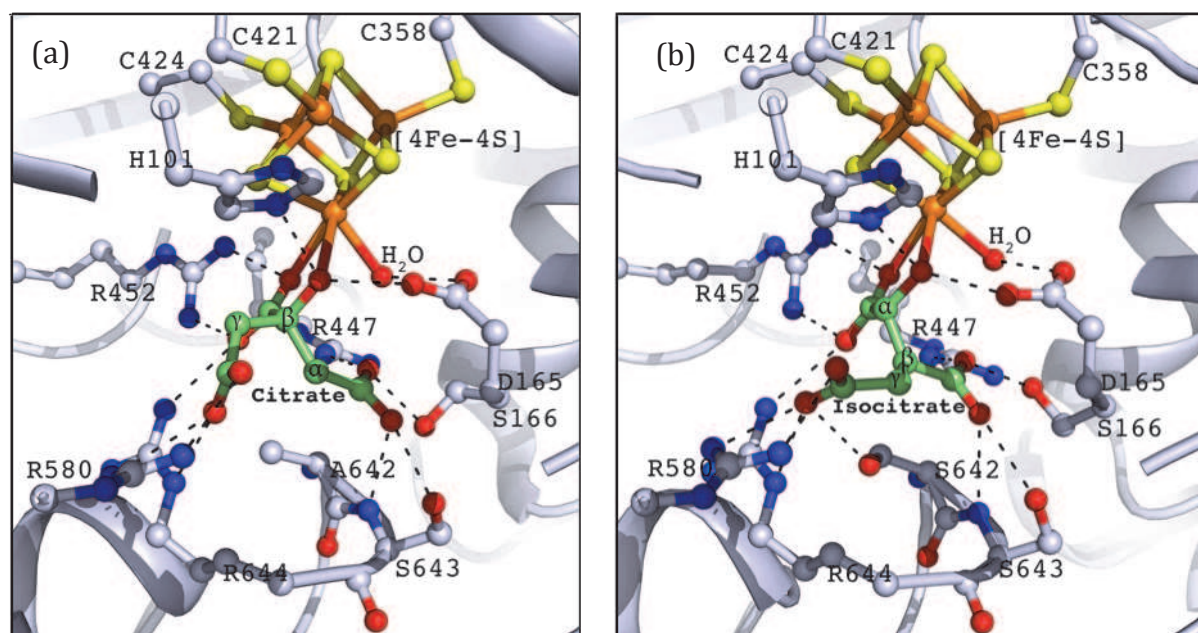
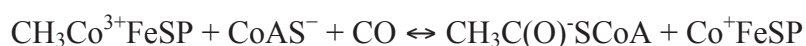


Figure I-4 Ligands citrate ou isocitrate liés à la m-aconitase.

(a) molécule de citrate liée à la m-aconitase (code PDB 1c96, (Lloyd, Lauble et al. 1999) ; (b) molécule d'isocitrate liée à la m-aconitase (code PDB 7acn, (Lauble, Kennedy et al. 1992). Les atomes de carbone des molécules de citrate et isocitrate sont en vert.

Dans le cas des protéines à radical SAM, le centre [4Fe-4S] possède aussi un atome de Fe libre qui coordine la *S*-Adénosyl-L-Méthionine (SAM) (Figure I-5 (a)). Dans ce cas le type de réaction catalysée est redox car le centre [4Fe-4S] va donner un électron à la SAM. Cette famille de protéines sera étudiée plus tard.

Un site pour la liaison à un substrat et son activation peut aussi être généré par la liaison d'un centre métallique sur un ion fer d'un centre FeS. L'acétyl-coenzymeA synthase (ACS) catalyse la réaction :



avec CoAS le coenzyme A et CoFeSP la protéine corrinoïde à centre FeS contenant du Co. Son site actif (« A-cluster », décrit précédemment) contient un centre [4Fe-4S] qui lie par un atome de Fe, un centre à 2 atomes de Ni. Il y a deux sites vacants sur un des atomes de Ni pour les substrats (Figure I-5 (b)). Dans l'hydrogénase à [Fe-Fe] le centre [4Fe-4S] lie un centre à 2 atomes de Fe, un site vacant pour un ligand est créé sur un des atomes de Fe, rendant possible la fixation d'hydrogène à ce site (décrit plus en détails dans la partie suivante).

Un site pour la liaison à un substrat et son activation peut aussi être généré par incorporation d'un autre métal que le Fe dans un centre FeS. La CO déshydrogénase (CODH),

catalyse la réaction réversible de réduction de CO_2 en CO . Son site actif (« C-cluster », décrit précédemment) contient une unité $[3\text{Fe-4S}]$ liée à un atome de fer et un atome de nickel. Il y a deux sites vacants sur l'atome de Ni pour les substrats (Figure I-5 (c)). Pour plus de détails, la revue de Fontecilla *et al.* (Fontecilla-Camps, Amara et al. 2009) décrit les protéines ACS et CODH.

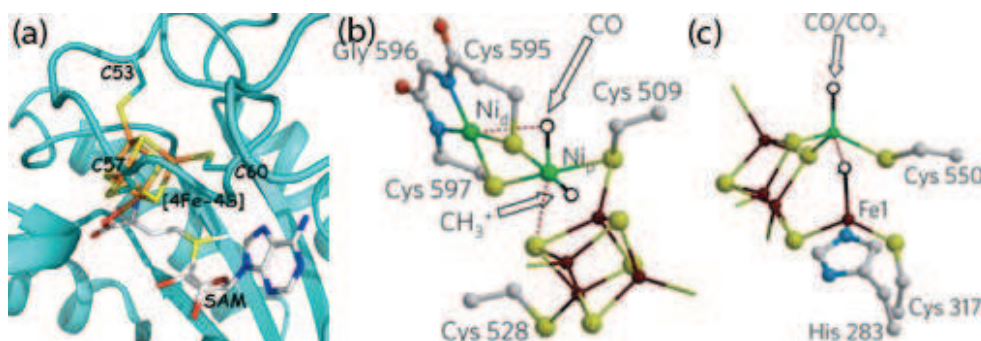


Figure I-5 Liaison à un substrat et son activation.

(a) protéine BioB (code PDB 1r30), (b) ACS, (c) CODH. Dans (b) et (c), les cercles vides indiquent les sites de liaison libres pour accueillir un substrat et les flèches indiquent les sites d'entrée/sortie des substrats. Les figures (b) et (c) sont adaptées de Fontecilla *et al.* (Fontecilla-Camps, Amara et al. 2009).

Les atomes de Fe sont représentés en rouge foncé ou orange, de soufre en jaune, de nickel en vert et d'azote en bleu, et d'oxygène en rouge.

I.1.2.3) La détection de conditions environnementales et régulation de l'expression de gènes

Il existe plusieurs exemples de centre FeS impliqués dans la régulation de la transcription et traduction de protéines.

La protéine IscR est un facteur de transcription régulateur de l'expression de plus de 40 gènes dans *E. coli* (Giel, Rodionov et al. 2006). Le rôle le plus connu de cette protéine est la régulation de l'opéron *isc* dont elle fait partie (Schwartz, Giel et al. 2001). Lorsqu'elle est chargée en centre $[2\text{Fe-2S}]$, IscR régule négativement la transcription par son propre promoteur situé avant le gène IscR. Lorsque la protéine IscR ne possède pas de centre FeS, elle active l'opéron *suf*. Cette protéine a donc un rôle de détection de conditions environnementales : la concentration faible en protéine IscR chargée reflète une efficacité réduite de la fabrication de centres FeS, et ceci est alors corrigé en synthétisant les protéines de l'opéron *isc* et *suf* (Giel, Rodionov et al. 2006).

La protéine FNR est un régulateur global d'*E. coli* qui contrôle l'expression de plus de 100 gènes en réponse à une limitation d'oxygène. La liaison à l'ADN pour l'activation de la

transcription de ces gènes ne se fait que lorsque la protéine est sous sa forme homodimérique. La formation du dimère n'a lieu que lors de la présence d'un centre [4Fe-4S] par monomère, lié par 4 cystéines. En présence d'oxygène, le centre est rapidement dégradé en centre [2Fe-2S] qui ne permet pas une dimérisation de la protéine. La dimérisation de la protéine se ferait donc après un changement conformationnel en présence du centre [4Fe-4S] (Kiley and Beinert 2003).

I.2) Les hydrogénases

Le métabolisme de l'hydrogène par les micro-organismes est fait par les hydrogénases. Il existe trois types d'hydrogénases : à [Fe-Fe], [Ni-Fe] et [Fe]. Les enzymes à [Fe-Fe] et [Ni-Fe] catalysent la réaction $H_2 \leftrightarrow 2H^+ + 2e^-$, tandis que l'hydrogénase à [Fe] catalyse la réaction d'hydrogénation de la 5,10-méthényl-tétrahydrométhanoptérine en 5,10-méthylène-tétrahydrométhanoptérine. Cette dernière ne se trouve que chez certains organismes archées méthanogènes ; l'hydrogénase à [Fe-Fe] se trouve chez les bactéries et quelques eucaryotes unicellulaires et l'hydrogénase à [Ni-Fe] chez les bactéries et les archées (Vignais and Billoud 2007).

Des résultats structuraux et spectroscopiques ont été publiés pour les 3 hydrogénases, révélant la singularité de leurs sites actifs métalliques (Fontecilla-Camps, Volbeda et al. 2007). Ces sites sont représentés dans la Figure I-6. Le site actif de l'hydrogénase à [Ni-Fe] est composé d'un centre contenant un atome de Ni et un atome de Fe. L'atome de Ni est coordiné par 4 cystéines dont 2 pontantes qui coordinent aussi l'atome de Fe. Ce dernier possède aussi deux ligands CN^- et un ligand CO.

Le site actif de l'hydrogénase à [Fe-Fe] (appelé « H-cluster ») est composé d'un centre [4Fe-4S] lié à un centre à deux atomes de Fe par une de ses cystéines. Chaque atome de fer est coordiné par un ligand CO et un ligand CN^- , et ces atomes de Fe sont reliés par un CO et les deux thiolates d'une petite molécule organique de formule proposée $^-S-CH_2-N(H)-CH_2-S^-$ (Nicolet, de Lacey et al. 2001).

Le site actif de l'hydrogénase à [Fe] est composé d'un cofacteur appelé FeGP constitué d'un atome de fer lié à deux ligands CO, l'azote d'un dérivé de 2-pyridone et un groupement acyl lié à ce dérivé (Hiromoto, Ataka et al. 2009). Ce cofacteur est relié par une cystéine de l'enzyme. Les 3 enzymes montrent un site vacant sur un atome de fer pour un ligand, qui pourrait être la molécule de H_2 .

Malgré le fait que les hydrogénases ne soient pas reliées phylogénétiquement, leurs sites actifs comportent des éléments similaires : une unité $\text{Fe}(\text{CO})_n$, avec $n=1$ pour l'hydrogénase à [Ni-Fe] et 2 pour les hydrogénases à [Fe-Fe] et [Fe]. Les hydrogénases à [Fe-Fe] et [Ni-Fe] possèdent aussi deux ligands CN^- .

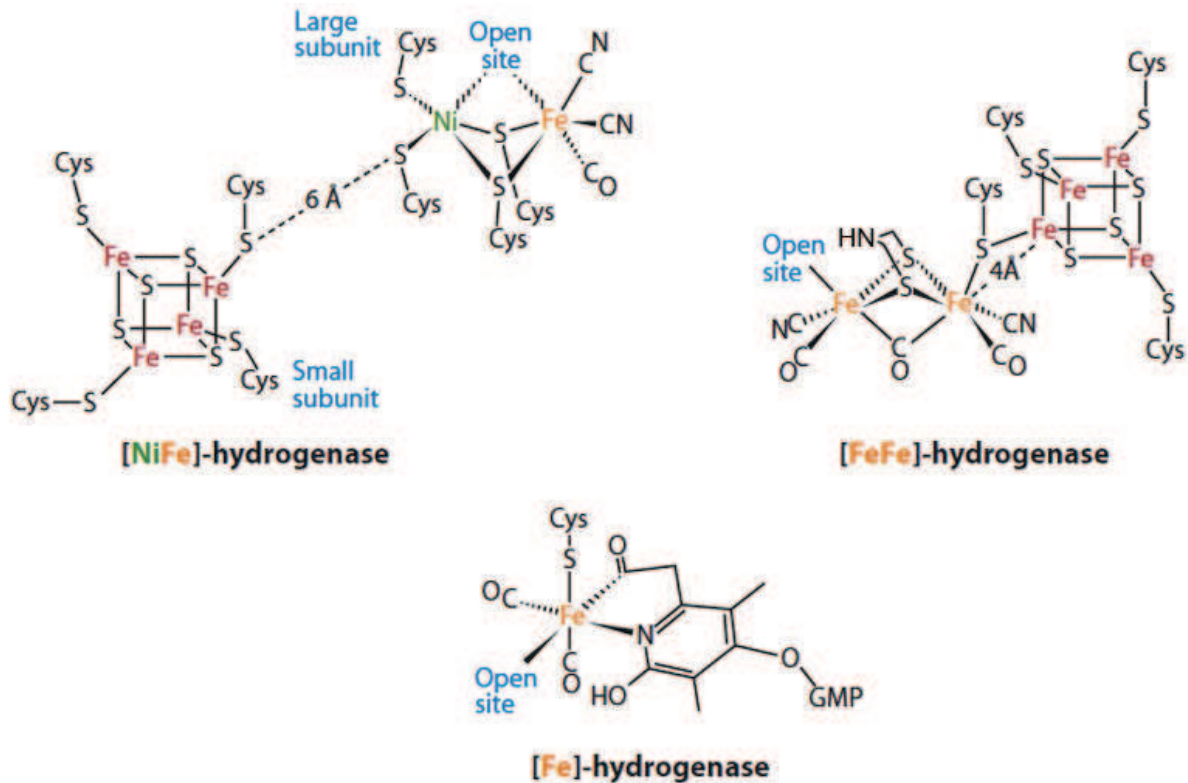


Figure I-6 Sites actifs des 3 types d'hydrogénase.

Figure tirée de (Thauer, Kaster et al. 2010). GMP : guanosyl monophosphate. Les sites vacants pour un ligand sont notés « Open site ».

Les centres FeS sont donc en général très variés et très utilisés, allant du plus « simple » atome de Fe de la rubrédoxine, puis centre $[2\text{Fe}-2\text{S}]$ aux centres FeS plus compliqués, comme ceux des hydrogénases.

II) Fabrication des centres FeS

Les centres FeS peuvent être synthétisés *in vitro* dans l'apo-protéine purifiée en utilisant des ions ferreux ou ferriques, du sulfure et un réducteur de type DTT. Cependant, les centres FeS les plus communs ($[2\text{Fe}-2\text{S}]$ et $[4\text{Fe}-4\text{S}]$) sont synthétisés par des machineries protéiques *in vivo*. Ceci peut être expliqué par le fait que les atomes de Fe et S libres sont toxiques pour la cellule.

II.1) Maturation de la nitrogénase et découverte des machineries ISC et SUF

Les nitrogénases sont des enzymes catalysant la formation d'ammoniaque à partir d'une molécule d'azote. Elles se trouvent, par exemple, dans *Azotobacter vinelandii*, un organisme diazotrophe. Il existe trois types de nitrogénase contenant du Mo, V ou Fe (Rubio and Ludden 2005), mais le type de nitrogénase le plus courant est celle qui contient du Mo. Elle est composée de deux protéines, la dinitrogénase et la dinitrogénase réductase. La dinitrogénase, un tétramère $\alpha_2\beta_2$, aussi appelée protéine à MoFe, contient deux paires de centres FeS complexes qui sont le « P-cluster » et le « FeMo-co ». La dinitrogénase réductase, aussi appelée protéine à Fe, est un dimère contenant un centre [4Fe-4S] et deux sites de liaison à Mg-ATP.

Les produits primaires de translation des gènes codant pour ces deux protéines sont inactifs. De nombreux gènes sont requis pour la biosynthèse et l'insertion des centres métalliques associés à la nitrogénase. L'identification de gènes spécifiques à la fixation d'azote, directement impliqués dans l'assemblage de centres FeS pour l'activation de la nitrogénase, a été faite par une analyse de gènes dont l'expression est coordonnée à celle de gènes des composants structuraux de la nitrogénase chez *Azotobacter vinelandii*. L'inactivation de gènes *nif* (*nif* pour **n**itrogen **f**ixation) engendre des défauts de maturation de la nitrogénase. Néanmoins, la délétion des gènes *nifS* ou *nifU* n'a pas abouti à une abolition mais seulement à une diminution de l'activité nitrogénase. Ils ne sont donc pas essentiels à son activation ((Jacobson, Brigle et al. 1989), (Jacobson, Cash et al. 1989)).

La fonction de la protéine NifS a été élucidée par Zheng *et al.* en 1993 (Zheng, White et al. 1993) : elle catalyse la formation de L-alanine et de soufre en utilisant de la L-cystéine. Cette enzyme utilise le phosphate de pyridoxal (PLP) comme cofacteur. Il se trouve dans chaque monomère de cette protéine dimérique. Elle a été nommée cystéine désulfurase. Durant la réaction, il se forme un persulfure sur un résidu de cystéine conservé que l'on appellera aussi cystéine active (Zheng, White et al. 1994) (Figure II-1).

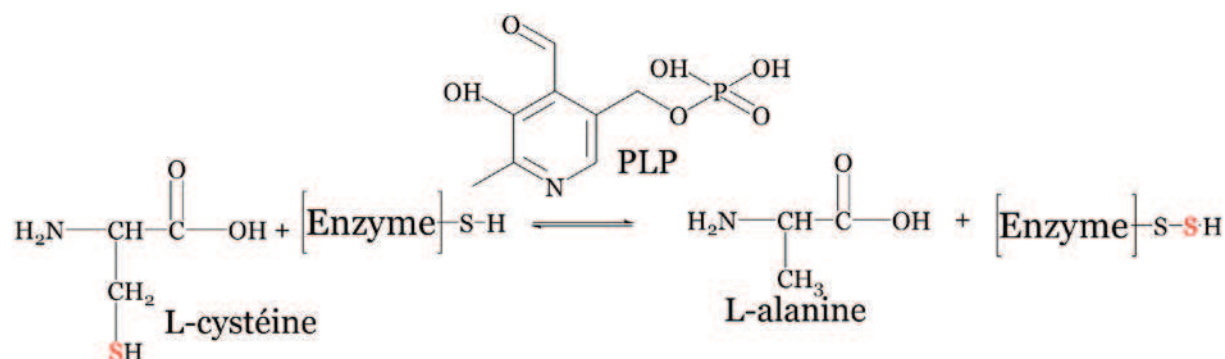


Figure II-1 Réaction catalysée par les cystéines désulfurases.

La cystéine désulfurase est nommée [Enzyme] et la chaîne latérale de sa cystéine active est représentée par « -SH ». « -S-SH » représente la cystéine active persulfurée.

NifS est capable d'assembler *in vitro* le centre [4Fe-4S] de la protéine à Fe de la nitrogénase (Zheng and Dean 1994). Elle aurait donc un rôle *in vivo* dans la mobilisation du soufre pour la fabrication de centres FeS de la nitrogénase.

Des cystéines désulfurases homologues à NifS ont été caractérisées : une en délétant le gène codant pour la protéine NifS d' *A. vinelandii* (Zheng, Cash et al. 1998) et une provenant d'un organisme non azotrophe, *E. coli* (Flint 1996). Elles sont capables de fabriquer des centres FeS autre que celui de la nitrogénase et ont été nommées IscS (isc pour iron-sulfur cluster) (ou « NifS-like »).

Le gène *iscS* d'*Av* se trouve à proximité de gènes appelés *iscU* et *iscA*, similaires à *nifU* et *orf6* dans les gènes *nif* (Zheng, Cash et al. 1998). Dans *E. coli*, à proximité des ces gènes se trouvent les gènes *hscB* et *hscA* (*hsc* pour *heat-shock-cognate*), codant pour des protéines similaires aux protéines chaperones dnaJ et dnaK (Lelivelt and Kawula 1995), ainsi qu'un gène *fdx*. Ces protéines chaperonnes pourraient assister la maturation de centres FeS de protéines. L'analyse des gènes d'*Av* en aval à *iscSUA* révèle la présence d'homologues des gènes *hscBA*, de *fdx* codant pour une protéine homologue de [2Fe-2S]-ferrédoxine. Un gène *orf3* se trouve à proximité et possède un homologue chez *E. coli*. L'agencement *iscSUA-hscBA-fdx* semble donc montrer une possible co-transcription de ces gènes (Zheng, Cash et al. 1998). La coexpression du cluster de gènes *orf1-orf2-iscSUA-hscBA-fdx-orf3* avec différentes ferrédoxines à centres [2Fe-2S] ou [4Fe-4S] conduit à l'obtention d'holoprotéines (Nakamura, Saeki et al. 1999). Afin d'identifier plus précisément les gènes impliqués dans la fabrication du centre FeS, la même expérience a été faite en délétant individuellement chaque gène du cluster. L'ensemble de gènes *orf2-iscSUA-hscBA-fdx-orf3* joue un rôle dans la fabrication des centres FeS car leur délétion amène à une diminution voire une élimination d'obtention

d'holoferrédoxines (Takahashi and Nakamura 1999). La délétion du gène *orf1* n'a pas modifié le taux d'assemblage d'agrégats FeS. Une souche d'*E. coli* délétée seulement du gène *iscS* a été produite afin de déterminer son rôle. L'activité des protéines contenant un centre FeS diminue considérablement en son absence, alors que les protéines n'en contenant pas ne voient pas leur activité modifiée (Schwartz, Djaman et al. 2000). La même étude a été faite mais en délétant aussi les autres gènes du cluster individuellement : les activités des protéines contenant un centre FeS ont chuté lors de la délétion de chacun des gènes du cluster, mis à part pour *orf2* (renommé *iscR*) et *orf3* (Tokumoto and Takahashi 2001). Les rôles d'IscR et ORF3 restaient donc à être élucidés, mais néanmoins, la machinerie de fabrication des centres FeS ISC était découverte.

Le rôle d'*iscR* a été trouvé en le délétant. La protéine IscR, chargée en centre [2Fe-2S] lors de sa purification, régule négativement la transcription de son propre promoteur situé avant le gène *IscR*. IscR régule l'opéron *iscRSUA* (Schwartz, Giel et al. 2001). Son rôle possible dans la régulation des gènes en aval *hscBA-fdx-orf3* n'a pas été étudié.

En mutant l'ensemble de l'opéron *isc* d'*E. coli* (*iscRSUA-hscBA-orf3*), une activité résiduelle d'enzymes à centres FeS était observée. D'autres gènes impliqués dans la fabrication de centres FeS ont été identifiés, ceux de l'opéron *SUF*, codant la machinerie *SUF* (**sulfur mobilization**). La délétion de cet opéron en présence de l'opéron *ISC* n'a pas montré d'effet sur l'activité de protéines à centres FeS (Takahashi and Tokumoto 2002), mais a montré son rôle dans des conditions de privation en fer (Outten, Djaman et al. 2004) ou de stress oxydant (Nachin, Loiseau et al. 2003).

II.2) La machinerie ISC

II.2.1) Exemples d'opérons codant pour les machineries bactériennes de fabrication des centres FeS

Trois machineries sont dédiées à la fabrication de centres FeS chez les bactéries, et les gènes codant pour ces machineries sont organisés en opérons (Figure II-2). Les gènes *nif* codent pour les protéines composant la nitrogénase ou celles impliquées dans sa maturation. Le système *NIF* se définit par les protéines impliquées dans la fabrication des centres FeS de cette dernière. Le système *ISC* est la machinerie primaire pour la biosynthèse de centres FeS. Le système *SUF* est opérationnel dans les bactéries subissant des conditions de stress oxydant ou de privation en fer.

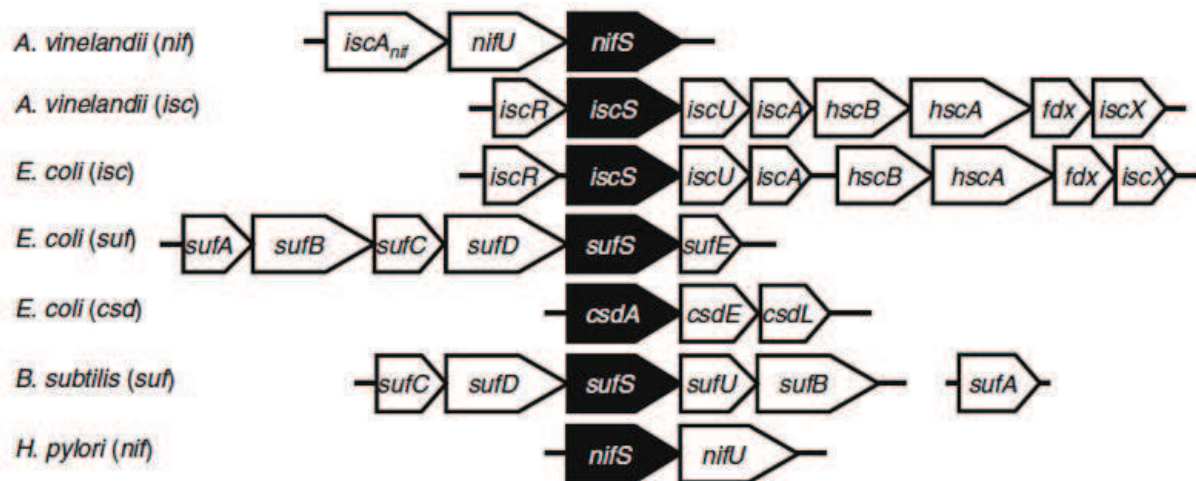


Figure II–2 Organisation des gènes dans des clusters de gènes *nif*, *isc*, *suf* et *csd*.

Figure tirée de la référence (Hidese, Mihara et al. 2011). Les machineries NIF, ISC et SUF sont codées par les clusters de gène *nif*, *isc* et *suf* respectivement.

Le principe de ces machineries est le même : il s'agit d'assembler un centre $[2\text{Fe-2S}]$ ou $[4\text{Fe-4S}]$ sur une protéine échafaudage, le soufre étant apporté par une cystéine désulfurase, et le fer par une autre protéine. Cette dernière semblerait être la protéine CyaY, homologue de la frataxine eucaryote (Py and Barras 2010). Le centre FeS est alors transféré à une apoprotéine.

La machinerie ISC est composée des protéines IscR, régulatrice de l'opéron *isc*; IscS, cystéine désulfurase ; IscU, protéine échafaudage ; IscA serait une protéine de transfert de centre FeS; protéines chaperonnes HscB et HscA ; une ferrédoxine et une protéine IscX dont le rôle n'est pas connu à ce jour. La protéine CyaY ne se trouve pas dans cet opéron.

II.2.2) Les protéines *IscS* et *IscU*

II.2.2.1) La protéine *IscS*

Les cystéines désulfurases se classent en deux groupes selon leur similarité de séquence. Les protéines des deux groupes se différencient nettement au niveau de 4 régions, incluant celle où se trouve la cystéine active conservée (Mihara, Kurihara et al. 1997). Pour cette dernière région, il s'agit de séquences similaires à SSGSACTS pour le groupe I et RXGHHCA pour le groupe II (Mihara and Esaki 2002). Le groupe I est représenté par les enzymes NifS, IscS, NFS et le groupe II par SufS (aussi appelée CsdB) et CsdA (aussi appelée CSD). IscS d'*E. coli* a 40% d'identité de séquence avec NifS d'*A. vinelandii*. CSD et CsdB ont seulement 24% d'identité avec NifS. La protéine CsdA est la troisième cystéine

désulfurase d'*E. coli* (cluster de gènes *csd* (Figure II–2), (Trotter, Vinella et al. 2009)). La cystéine désulfurase des eucaryotes est nommée NFS ou Nfs1. L'alignement multiple de séquences de Nfs1, NifS et IscS du groupe I montre clairement la similarité entre les différentes séquences. Le motif GSAC contenant la cystéine active est conservé (Figure II–3).

```

NifS_Azotobacter_vinelandii
NifS_Klebsiella_pneumoniae
Nfs1_Homo_sapiens
Nfs1_Mus_musculus
IscS_Caenorhabditis_elegans
Nfs1_Saccharomyces_cerevisiae
IscS_Escherichia_coli
IscS_Haemophilus_influenzae
IscS_Azotobacter_vinelandii
IscS_Neisseria_meningitidis
IscS_Archaeoglobus_fulgidus
IscS_Aquifex_aeolicus

```

```

NifS_Azotobacter_vinelandii
NifS_Klebsiella_pneumoniae
Nfs1_Homo_sapiens
Nfs1_Mus_musculus
IscS_Caenorhabditis_elegans
Nfs1_Saccharomyces_cerevisiae
IscS_Escherichia_coli
IscS_Haemophilus_influenzae
IscS_Azotobacter_vinelandii
IscS_Neisseria_meningitidis
IscS_Archaeoglobus_fulgidus
IscS_Aquifex_aeolicus

```

```

NifS_Azotobacter_vinelandii
NifS_Klebsiella_pneumoniae
Nfs1_Homo_sapiens
Nfs1_Mus_musculus
IscS_Caenorhabditis_elegans
Nfs1_Saccharomyces_cerevisiae
IscS_Escherichia_coli
IscS_Haemophilus_influenzae
IscS_Azotobacter_vinelandii
IscS_Neisseria_meningitidis
IscS_Archaeoglobus_fulgidus
IscS_Aquifex_aeolicus

```

```

NifS_Azotobacter_vinelandii
NifS_Klebsiella_pneumoniae
Nfs1_Homo_sapiens
Nfs1_Mus_musculus
IscS_Caenorhabditis_elegans
Nfs1_Saccharomyces_cerevisiae
IscS_Escherichia_coli
IscS_Haemophilus_influenzae
IscS_Azotobacter_vinelandii
IscS_Neisseria_meningitidis
IscS_Archaeoglobus_fulgidus
IscS_Aquifex_aeolicus

```

```

NifS_Azotobacter_vinelandii
NifS_Klebsiella_pneumoniae
Nfs1_Homo_sapiens
Nfs1_Mus_musculus
IscS_Caenorhabditis_elegans
Nfs1_Saccharomyces_cerevisiae
IscS_Escherichia_coli
IscS_Haemophilus_influenzae
IscS_Azotobacter_vinelandii
IscS_Neisseria_meningitidis
IscS_Archaeoglobus_fulgidus
IscS_Aquifex_aeolicus

```

```

NifS_Azotobacter_vinelandii
NifS_Klebsiella_pneumoniae
Nfs1_Homo_sapiens
Nfs1_Mus_musculus
IscS_Caenorhabditis_elegans
Nfs1_Saccharomyces_cerevisiae
IscS_Escherichia_coli
IscS_Haemophilus_influenzae
IscS_Azotobacter_vinelandii
IscS_Neisseria_meningitidis
IscS_Archaeoglobus_fulgidus
IscS_Aquifex_aeolicus

```



Figure II-3 Alignement multiple de séquences de cystéines désulfurases du groupe I.

Alignement multiple des séquences d'IscS d'*Escherichia coli* (gi|49176235), IscS d'*Haemophilus influenzae* (gi|30995370), IscS d'*Azotobacter vinelandii* (gi|226946078), IscS de *Neisseria meningitidis* (gi|304387293), Nfs1 d'*Homo sapiens* (gi|189055336), Nfs1 de *Mus musculus* (gi|161016824), IscS de *Caenorhabditis elegans* (gi|25143064), Nfs1 de *Saccharomyces cerevisiae* (gi|10383773), IscS d'*Archaeoglobus fulgidus* (gi|11498172), IscS d'*Aquifex aeolicus* (gi|15606336), NifS d'*Azotobacter vinelandii* (gi|226942328), NifS de *Klebsiella pneumoniae* (gi|206577044). L'étoile violette indique la position de la cystéine active catalytique.

Les différentes cystéine désulfurases sont structuralement homologues (Figure II-4 (a)). La différence au niveau de la région de la cystéine conservée se voit aussi au niveau des structures. La cystéine active de la protéine IscS d'*E. coli* (Cupp-Vickery, Urbina et al. 2003), code PDB 1p3w) n'est pas visible dans sa structure. Elle se situe dans une boucle non visible donc flexible (Figure II-4 (b)). Dans le cas de la protéine SufS de *Synechocystis* sp. PCC 6803 (Tirupati, Vey et al. 2004), code PDB 1t3i) la cystéine active est visible dans les cristaux, et elle se trouve proche d'une hélice (Figure II-4 (b)).

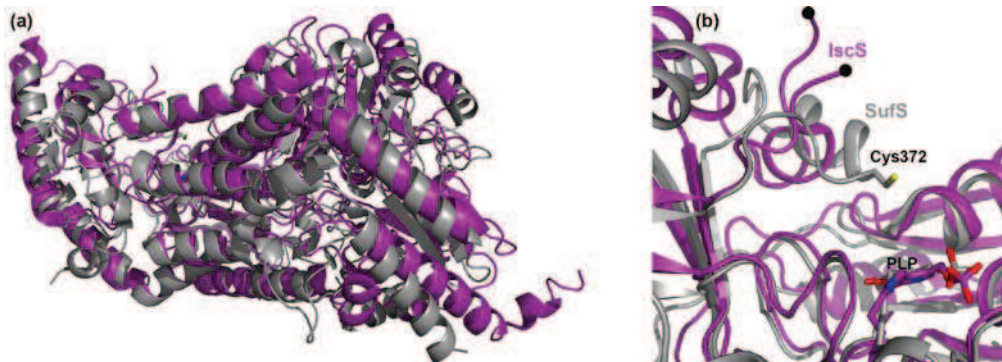


Figure II-4 Superposition de structures de cystéines désulfurases.

(a) Superposition des structures de *EcIscS* (code PDB 1P3W) en violet et *SufS* de *Synechocystis* sp. PCC 6803 (code PDB 1t3i) en gris, avec, à droite, (b) un zoom sur la région de la cystéine active. Les cercles noirs dans (b) pour *IscS* indiquent les derniers résidus visibles de chaque extrémité de la boucle contenant la cystéine active ; la cystéine active Cys372 de *SufS* est représentée en bâtons.

II.2.2.2) La protéine IscU

La fonction exacte de la protéine IscU a été élucidée en 2000. Il s'agit d'une protéine échafaudage sur laquelle se fabrique un centre [2Fe-2S] ou [4Fe-4S], IscS étant le donneur de soufre ((Agar, Zheng et al. 2000), (Agar, Krebs et al. 2000)), et le fer étant probablement apporté par la protéine CyaY. Il existe deux types de protéines échafaudage : le type U et le type NFU.

Les protéines échafaudage de type U sont caractérisées par un motif Cx₂₄₋₂₆Cx₄₂₋₄₃C dans leur séquence, qui est impliqué dans la ligation du cluster. Les membres de cette classe sont les protéines NifU (machinerie NIF), IscU (machinerie ISC), ISU (homologue eucaryote de IscU) et SufU (machinerie SUF) (Fontecave and Ollagnier-De-Choudens 2008). Néanmoins SufU (i) possède 18 à 21 acides aminés en plus entre la 2^{ème} et 3^{ème} cystéine ; (ii) contient une lysine précédant la 3^{ème} cystéine au lieu d'une histidine dans le cas d'IscU/NifU/ISU ; (iii) ne contient pas le motif LPPVK présent dans IscU impliqué dans la liaison IscU/HscA (Johnson, Dean et al. 2005). Toutefois, l'alignement multiple de séquences des protéines IscU/NifU/ISU montre clairement leurs similarités (Figure II-5).

La protéine NifU possède 9 cystéines conservées réparties dans 3 domaines de la protéine. Le motif de type U se situe dans le domaine N-terminal de la protéine. Le motif de type NFU Cx₂C est représenté par le domaine C-terminal de la protéine NifU.

Comme on peut le voir sur la Figure II-6, qui représente la structure d'*Ec*IscU, les 3 cystéines conservées du motif de type U de IscU (Cys33, Cys58, Cys102) sont à proximité de l'acide aspartique (Asp35) conservé mentionné dans le paragraphe suivant. L'histidine précédant la 3^{ème} cystéine est l'His101. Le motif LPPVK se trouve à la surface de la protéine et à proximité.

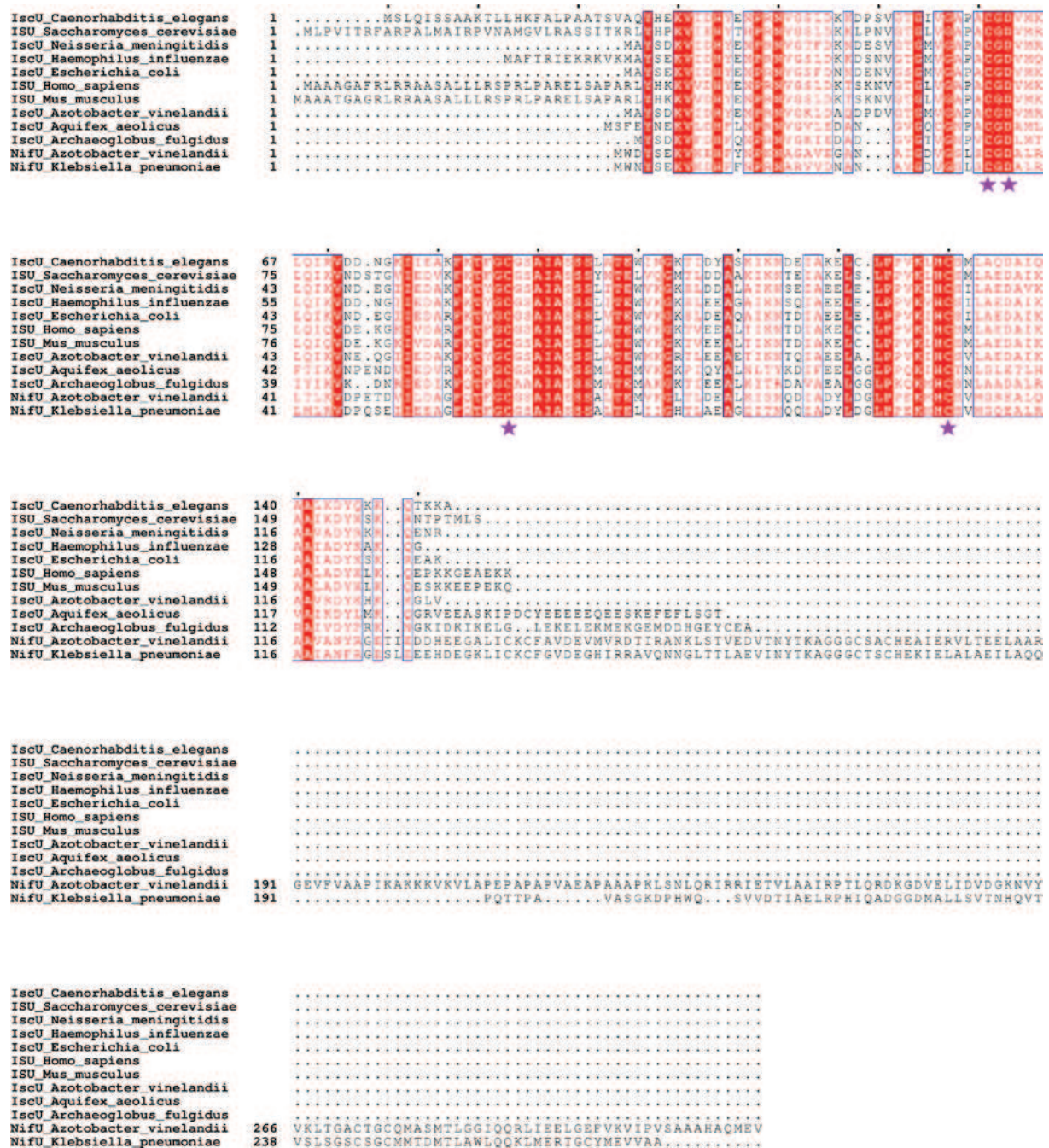


Figure II-5 Alignement multiple de séquences de protéines IscU et d'homologues.

Alignement multiple des séquences d'IscU d'*Escherichia coli* (gi|16130454), IScU d'*Haemophilus influenzae* (gi|30995369), IScU d'*Azotobacter vinelandii* (gi|226946077), IScU de *Neisseria meningitidis* (gi|304387291), ISU d'*Homo sapiens* (gi|56699456), ISU de *Mus musculus* (gi|21313484), IScU de *Caenorhabditis elegans* (gi|17543474), ISU de *Saccharomyces cerevisiae* (gi|6325122), IScU d'*Archaeoglobus fulgidus* (gi|11497802), IScU d'*Aquifex aeolicus* (gi|15606229), NifU d'*Azotobacter vinelandii* (gi|226942327), NifU de *Klebsiella pneumoniae* (gi|128319). Les étoiles violettes indiquent les positions des cystéines et l'acide aspartique conservés.

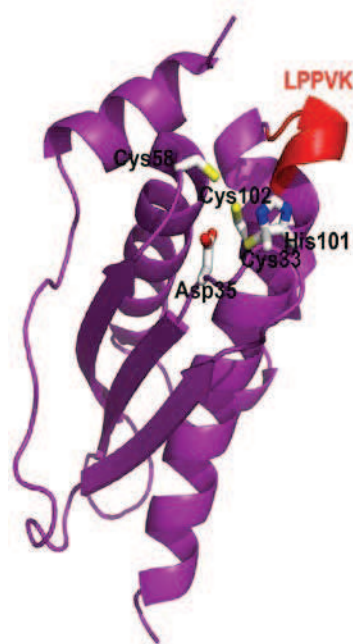


Figure II–6 Structure d’*Ec*IscU.

Structure d’*Ec*IscU (code PDB 3lv1, (Shi, Proteau et al. 2010)). Les résidus du motif LPPVK sont colorés en rouge.

II.2.2.3) Assemblage de centres [2Fe-2S] ou [4Fe-4S] sur le domaine N-terminal de NifU et IscU



Figure II–7 Domaines de NifU.

Domaines N-terminal, central et C-terminal d’*Av*NifU avec les cystéines de chaque domaine.

NifU est une protéine homodimérique modulaire. Elle est composée de domaines N-terminal, central et C-terminal (Figure II–7). Le domaine central comporte un centre redox-actif [2Fe-2S] stable et résistant aux chélateurs de fer (Fu, Jack et al. 1994). Il est coordonné par 4 cystéines conservées (Agar, Yuvaniyama et al. 2000) et est homologue à une ferrédoxine (Smith, Jameson et al. 2005). Le domaine N-terminal, comportant un motif de type U, est un site d’assemblage de centres [2Fe-2S] (Yuvaniyama, Agar et al. 2000), mais non stables. Une forme altérée de la protéine permet d’obtenir un centre [2Fe-2S] stable dès l’isolement de la protéine : il s’agit de la mutation de l’acide aspartique conservé proche de la première cystéine conservée (Figure II–5) en alanine (Yuvaniyama, Agar et al. 2000). Ce domaine est aussi capable d’assembler des centres [4Fe-4S]. Il se forme séquentiellement un puis deux centres [2Fe-2S] et un centre [4Fe-4S] par dimère. La protéine est capable de le transférer à la Fe protéine de la nitrogénase. La mutation de l’acide aspartique en alanine empêche le transfert du cluster et de plus modifie les paramètres Mössbauer des centres [2Fe-2S] et [4Fe-4S] (Smith, Jameson et al. 2005).

La protéine IscU est aussi capable d'assembler des centres [2Fe-2S] et [4Fe-4S]. En mettant en contact les protéines IscS et IscU dans un rapport équimolaire, l'analyse par colonne de chromatographie d'exclusion stérique montre que les protéines interagissent entre elles sous forme d'hétérotétramère $\alpha_2\beta_2$ (Agar, Zheng et al. 2000). Comme pour le domaine N-terminal de NifU, en ce qui concerne la formation de centre [4Fe-4S], elle se fait séquentiellement et lentement par le passage d'un puis deux centres [2Fe-2S] et un centre [4Fe-4S] par dimère (Agar, Krebs et al. 2000). Lorsqu'on ajoute un équivalent de dithionite ou la fdx de l'opéron *isc*, le passage de la forme 2 centres [2Fe-2S] à un centre [4Fe-4S] par dimère, se fait très rapidement (Chandramouli, Unciuleac et al. 2007). La fdx de l'opéron *isc* pourrait donc avoir le rôle physiologique de donneur d'électron afin de promouvoir le couplage rapide de deux centres [2Fe-2S] et centre [4Fe-4S]. L'exposition à l'oxygène du centre [4Fe-4S] formé amène à l'obtention d'un centre [2Fe-2S] par dimère.

Les protéines chaperonnes HscB et HscA présentes dans l'opéron *isc* ont un rôle stimulant dans le transfert du centre [2Fe-2S] d'IscU à une apoferrédoxine. Le site de reconnaissance de la protéine HscA sur IscU est le motif LPPVK (Chandramouli and Johnson 2006). La protéine IscU est aussi capable de transférer son centre [4Fe-4S] à une apo-aconitase. La mutation de l'acide aspartique conservé en alanine empêche ce transfert. Les protéines chaperonnes n'ont aucun effet sur l'augmentation de la rapidité de ce transfert (Unciuleac, Chandramouli et al. 2007).

La protéine IscA aussi présente dans l'opéron *isc* serait une protéine de transfert de centre FeS, aussi appelée A-type carrier (en anglais) (abrégié ATC). Les ATCs possèdent un motif conservé Cx₆₃₋₆₅CxC. Elles sont proposées être des protéines de transfert de centres FeS, ces derniers leur ayant été transférés après avoir été fabriqués sur des protéines échafaudage de type U ou NFU. Elles pourraient alors transférer le centre à une apoprotéine cible (Py and Barras 2010). Elles seraient en quelque sorte des protéines de stockage de centres FeS.

II.3) Assemblage des centres FeS chez les eucaryotes

Pour ce qui est de la fabrication des centres FeS chez les eucaryotes, la mitochondrie possède des composants homologues à ceux de la machinerie ISC bactérienne et les plastides possèdent des composés homologues à ceux de la machinerie SUF. L'assemblage des centres FeS dans le cytosol et dans le noyau des eucaryotes requière l'assistance de la machinerie ISC de la mitochondrie et d'un système d'export de la mitochondrie vers le cytosol. Il se trouve en

plus une machinerie appelée CIA (« cytosolic iron-sulfur cluster assembly ») dans le cytosol (Py and Barras 2010). Les composants des machineries bactériennes et les composants possédant la même fonction dans les machineries eucaryotes sont indiquées dans le Tableau II-1 (Lill 2009).

Table 1 Compilation of central components involved in Fe-S-protein biogenesis					
Function	Bacterial NIF	Bacterial Suf	Bacterial ISC	Mitochondrial ISC	Eukaryotic CIA
Cysteine desulphurase, sulphur donor	NifS	SufS-SufE	IscS	Nfs1-Isc11	Mitochondrial Nfs1-Isc11
U-type scaffold for Fe-S-cluster assembly	NifU (amino-terminal domain)	SufU	IscU	IscU	-
A-type scaffold for Fe-S-cluster assembly	IscA ^{Nif}	SufA	IscA, ErpA	IscA, Isa2, Iba57?	-
NFU-type scaffold for Fe-S-cluster assembly	NifU (carboxy-terminal domain)	-	NfuA	Nfu1 (also present in plastids)	-
Scaffold for Fe-S-cluster assembly	-	SufB	-	-	-
P-loop NTPase scaffold for Fe-S-cluster assembly	-	-	-	Isc1?	Cfd1-Nbp35
Electron transfer	NifU (middle domain)	-	Fdx	Yah1-Arh1	-
Iron donor	-	-	CyaY	Yfh1 (frataxin)	-
Transfer of Fe-S cluster from scaffold to target proteins	-	SufC?	HscA, HscB	Ssq1, Jac1, Mge1, Grx5	Nar1, Cia1
P-loop NTPase, phosphate-binding-loop nucleoside triphosphatase.					

Tableau II-1 Composants impliqués dans la biogenèse des centres FeS.

Figure tirée de Lill (Lill 2009).

Chez l'humain, une déficience dans l'un des éléments de fabrication des centres FeS ou de certaines protéines contenant des centres FeS provoque des maladies (Tableau II-2).

Table 2 Diseases related to Fe-S proteins and their biogenesis			
Human protein	Yeast homologue	Function	Associated disease
Biogenesis components			
Mitoferrin	Mrs3, Mrs4	Mitochondrial carrier; putative iron transporter	Erythropoietic protoporphyria ⁹⁰
ISCU	IscU	Mitochondrial scaffold	Myopathy with exercise intolerance ^{91,92}
Frataxin	Yfh1	Iron donor for IscU scaffold	Friedreich's ataxia ⁹³
GLRX5	Grx5	Monothiol glutaredoxin; transfer of Fe-S clusters from IscU to apoproteins	Microcytic anaemia ⁹⁴
ADR	Arh1	Adrenodoxin reductase; electron transfer from NADH to adrenodoxin; also involved in steroid hormone biosynthesis	Tumour suppressor
ABC7	Atm1	ABC transporter, ISC export machinery component	X-linked sideroblastic anaemia and cerebellar ataxia ⁹⁵
Fe-S proteins			
Complex I	n.p.	Complex I of respiratory chain (NADH:ubiquinone oxidoreductase)	Various mitochondrial diseases such as LHON, MELAS, Leigh syndrome
Complex II	Complex II (Sdh1, Sdh2, Sdh3, Sdh4)	Complex II of respiratory chain (succinate dehydrogenase)	Tumour suppressor
XPD	Rad3	Fe-S protein involved in nucleotide excision repair	Xeroderma pigmentosum, Cockayne syndrome, trichothiodystrophy
FANCI	Rad3	Fe-S protein involved in nucleotide excision repair	Fanconi anaemia
MUTYH	Ntg2	Fe-S protein (glycosylase) involved in DNA repair	Colon cancer
n.p., not present in <i>S. cerevisiae</i> .			

Tableau II-2 Maladies associées à aux protéines à centres FeS et leur biogenèse.

Figure tirée de Lill (Lill 2009).

L'assemblage de centres FeS est donc un mécanisme complexe, et l'élucidation de ce mécanisme n'est toujours pas complète.

II.4) Identification d'une machinerie de maturation de l'hydrogénase à [Fe]

L'hydrogénase à [Fe], aussi appelée Hmd (ou HmdA : H₂ forming méthyle-H₄-methanopterin déhydrogénase), présente uniquement chez certains microorganismes archées méthanogènes, catalyse la réaction d'hydrogénation du 5,10-méthényl-tétrahydrométhanoptérine en 5,10-méthylène-tétrahydrométhanoptérine (Figure II-8) (Schleucher, Griesinger et al. 1994). Cette réaction fait partie du processus de méthanogénèse, c'est-à-dire la production de méthane à partir de dioxyde de carbone. Cette même réaction peut aussi être catalysée par la F₄₂₀H₂-dépendante méthylènetétrahydrométhanoptérine déshydrogénase (Mtd) (Hendrickson and Leigh 2008).

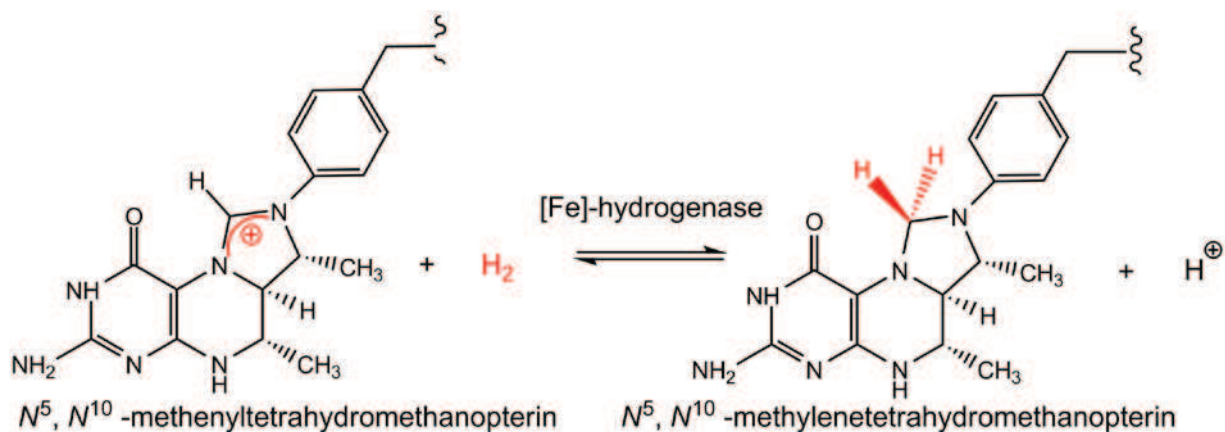


Figure II-8 Réaction catalysée par Hmd.

Hmd contient un « cofacteur FeGP » constitué d'un atome de fer lié à deux ligands CO, l'azote d'un dérivé de 2-pyridone, un groupement acyl lié à ce dérivé et un ligand inconnu (Figure II-9) (Hiromoto, Ataka et al. 2009). Le cofacteur est relié à la protéine par un résidu cystéine et l'activité de la protéine est dépendante de sa présence. Il est possible de l'extraire d'une protéine Hmd purifiée de son organisme hôte, pour ensuite l'insérer dans une protéine surexprimée dans *E. coli* (Buurman, Shima et al. 2000). Ceci suggère que le cofacteur serait synthétisé *in vivo* hors de la protéine et ensuite inséré dans cette dernière.

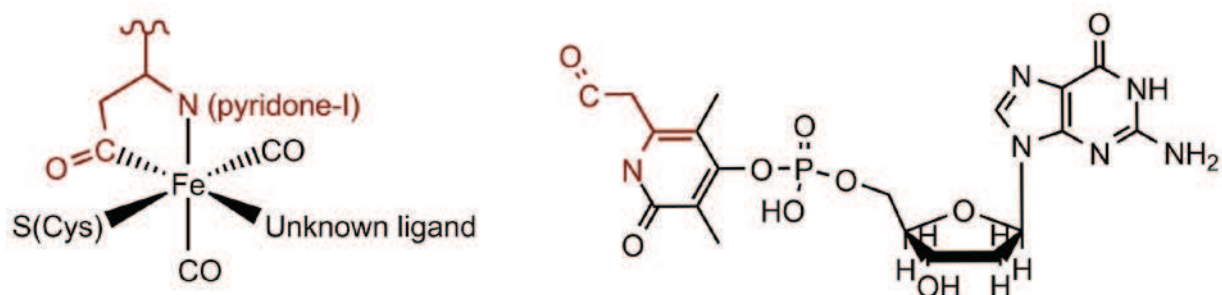


Figure II–9 Cofacteur FeGP d’Hmd.

La composition du cofacteur FeGP a été obtenue en utilisant des méthodes spectroscopiques diverses telles que le Mössbauer ou la RMN. L’ascension vers la détermination de la structure de ce cofacteur est détaillée dans la revue de Shima (Shima and Ermler 2011). La géométrie détaillée de ce cofacteur a été obtenue par la résolution de structures d’Hmd (Shima and Ermler 2011). Une de ces structures est représentée dans la Figure II–10.

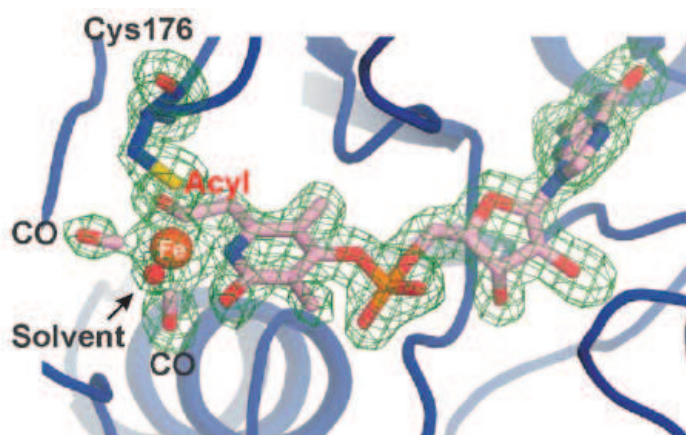


Figure II–10 Site actif d’Hmd.

Figure tirée de Shima et Ermler (Shima and Ermler 2011).

Les sites actifs des hydrogénases à [Fe-Fe] et [Ni-Fe] présentés plus haut sont assemblés par des machineries protéiques ((Fontecilla-Camps, Volbeda et al. 2007), (Boeck, King et al. 2006)). Il en est de même pour la nitrogénase (Rubio and Ludden 2005). Les gènes des protéines participant à l’assemblage de ces sites se situent en général à proximité du gène codant pour l’apoprotéine. Peu de travaux ont été réalisés sur le mécanisme d’assemblage du site actif d’Hmd. Néanmoins on peut supposer qu’il existe aussi une machinerie pour la biosynthèse et l’insertion de son cofacteur, d’autant plus que ce dernier peut être inséré *in vitro* dans une protéine apo.

Le Dr Yvain Nicolet, de notre équipe, avait découvert en 2005, en recherchant des protéines similaires à BioB, qu'une protéine à radical SAM (PRS) (famille présentée dans le paragraphe suivant) se trouvait à proximité du gène Hmd, et uniquement dans les organismes possédant Hmd. Cette protéine a été nommée HmdB. Elle a été caractérisée en 2010 par McGlynn *et al.* (McGlynn, Boyd et al. 2010).

Il y a deux PRS dans la machinerie de l'hydrogénase à fer-fer (présence des PRS HydE et HydG) (Mulder, Shepard et al. 2011) et de la nitrogénase (présence de la PRS NifB) (Rubio and Ludden 2008). Nous avons donc émis l'hypothèse que la protéine HmdB serait impliquée dans la biosynthèse du cofacteur FeGP de Hmd. Nous avons alors cherché à analyser les gènes voisins à Hmd et HmdB. Nous avons pour ça fait une recherche de séquence de Hmd de *Methanopyrus kandleri* (locus primaire MK0013) contre une base de données de séquences de protéines (basic local alignment search tool (BLAST) de protéine, (Altschul, Gish et al. 1990)). Comme il était attendu, les meilleurs scores d'alignement de séquence sont attribués à des protéines issues d'organismes méthanogènes (environ 55 % d'identité de séquence). Néanmoins, une protéine issue de l'organisme *Desulfurobacterium thermolithotrophum* possède 28% d'identité de séquence. *D. thermolithotrophum* est une bactérie réductrice du soufre mais il n'a pas été observé qu'elle était un archée méthanogène (L'Haridon, Cilia et al. 1998). Il est donc très étonnant de retrouver des homologues de ces gènes dans cet organisme.

Nous avons comparé les séquences des protéines voisines dans tous ces organismes. Des gènes voisins à Hmd codent pour des séquences protéiques similaires à travers ces organismes. Les gènes en question ont été représentés d'une même couleur dans la Figure II-11. Ils ne sont pas toujours agencés de la même manière dans l'ADN, et ne sont pas forcément à proximité. Nous avons cherché à identifier les protéines codées en utilisant l'outil de recherche de domaines protéiques conservés (CDD, (Marchler-Bauer, Lu et al. 2011)). Le gène MK0010 code pour une protéine homologue aux activateurs d'ubiquitine. Le gène MK0012 n'a pas de fonction reconnaissable. Le gène MK0014 code pour une protéine possédant un domaine de liaison au NADP de type Rossmann (Thauer, Kaster et al. 2010). Le gène MK0015 code pour un homologue de NIF3 qui interagit avec l'activateur transcriptionnel NGG1p chez la levure. Le gène MK0016 code pour HmdB. Le gène MK0017 code pour une protéine homologue à la fibrillarine. Le gène MK0018 code pour une protéine dont la structure a été résolue pour l'organisme *M. thermoautotrophicum*. Il s'agit d'un homodimère possédant 3 phosphates liés et partageant des similarités structurales avec les pyrophosphatases (Thauer, Kaster et al. 2010).

En plus de posséder un homologue de la protéine Hmd, *Desulfurobacterium thermolithotrophum* possède des homologues des gènes décrits ci-dessus, mis à part MK0018. Dans cet organisme le gène Dester_1507 possède en même temps des homologues des gènes MK0018 et MK0014.



Figure II–11 Gènes adjacents au gène codant pour la protéine Hmd.

Les gènes codants pour une protéine homologue sont de même couleur.

Le rôle dans la maturation de Hmd des protéines codées par les gènes à proximité du gène codant pour Hmd n'est pas clair. La simple analyse de gènes ne permet pas d'élucider le mécanisme de maturation mais a permis d'identifier la protéine HmdB. Afin d'élucider en partie ce mécanisme de maturation de la protéine Hmd, nous avons décidé d'étudier la protéine à radical SAM HmdB de *Methanopyrus kandleri*.

III) Les protéines à radical SAM

La superfamille des protéines à radical SAM (PRS) a été découverte par Sofia *et al.* en 2001 (Sofia, Chen et al. 2001). Ces protéines génèrent un radical 5'-désoxyadénosyl à partir de la SAM. Il s'agit d'une étape d'initiation radicalaire (Figure III-1). Le radical va alors arracher un atome d'hydrogène spécifique d'un substrat.

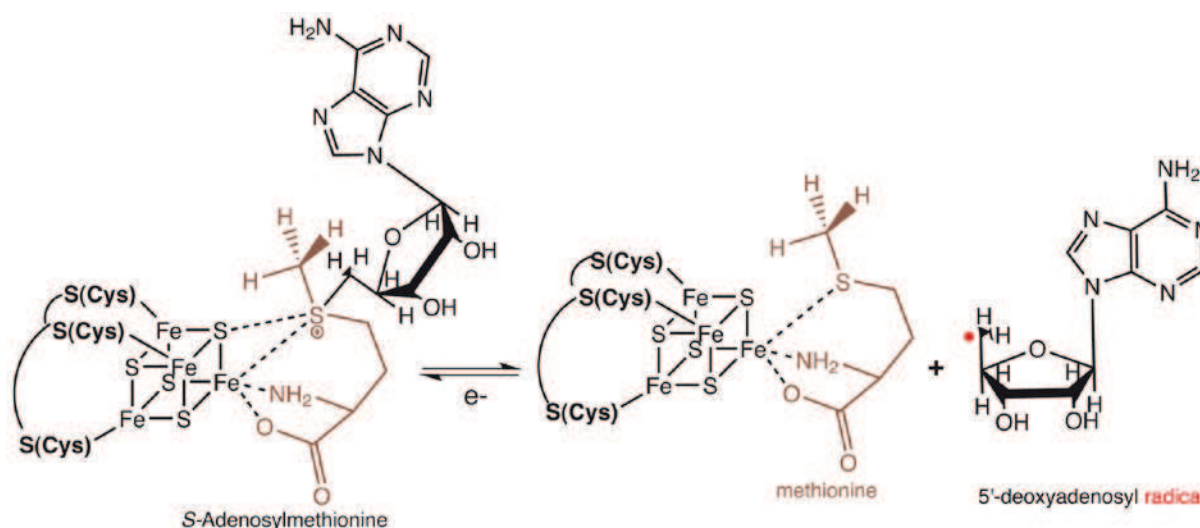


Figure III-1 Réaction d'initiation radicalaire des protéines à radical SAM.

Les substrats sont variés, chaque PRS ayant une spécificité fonctionnelle. Les réactions d'élimination, réarrangement, méthylation, activation d'enzyme, modification d'ARN, biosynthèse de cofacteur, oxydation, réparation d'ADN, insertion de soufre sont catalysées par ces enzymes (Marsh, Patterson et al. 2010).

III.1) Le cœur ($\beta\alpha$)₆ des PRS

Sur la base d'un alignement structural (Nicolet and Drennan 2004), on peut voir que la plupart des PRS possèdent le même cœur ($\beta\alpha$)₆ arrangé en $\frac{3}{4}$ de tonneau triose phosphate isomérase (TIM), responsable de la génération du radical. Le tonneau TIM se caractérise par un motif β/α se répétant 8 fois, les brins β formant un feuillet parallèle en forme de tonneau, et les hélices α entourant ce tonneau. A ce cœur s'ajoutent des éléments structuraux supplémentaires spécifiques à chaque protéine, afin de s'accommoder au type de substrat. Plusieurs structures de PRS ont été résolues à ce jour : celles de l'enzyme d'activation de la pyruvate formate lyase (PFL-AE) (Vey, Yang et al. 2008), coproporphyrinogène III oxydase

indépendante de l'oxygène (HemN) (Layer, Moser et al. 2003), biotine synthase (BioB) (Berkovitch, Nicolet et al. 2004), protéine de la biosynthèse du cofacteur à molybdène (MoaA) (Hanzelmann and Schindelin 2006), méthylthiotransférase RimO (Arragain, Garcia-Serres et al. 2010), méthyltransférase RlmN (Boal, Grove et al. 2011), lysine 2,3-aminomutase (LAM) (Lepore, Ruzicka et al. 2005), protéine de la biosynthèse des bases wye (TYW1) (Suzuki, Noma et al. 2007), maturase de l'hydrogénase à [Fe-Fe] (HydE) (Nicolet, Rubach et al. 2008). Elles possèdent toutes des structures différentes comme on peut le voir sur la Figure III-2, où sont représentées quelques structures, mais ont néanmoins un cœur commun ($\beta\alpha$)₆ (Figure III-3). La protéine RlmN possède une structure similaire à la PFL-AE, et RimO une structure similaire à celle de MoaA. RimO possède 3 domaines dont un domaine radical-SAM.

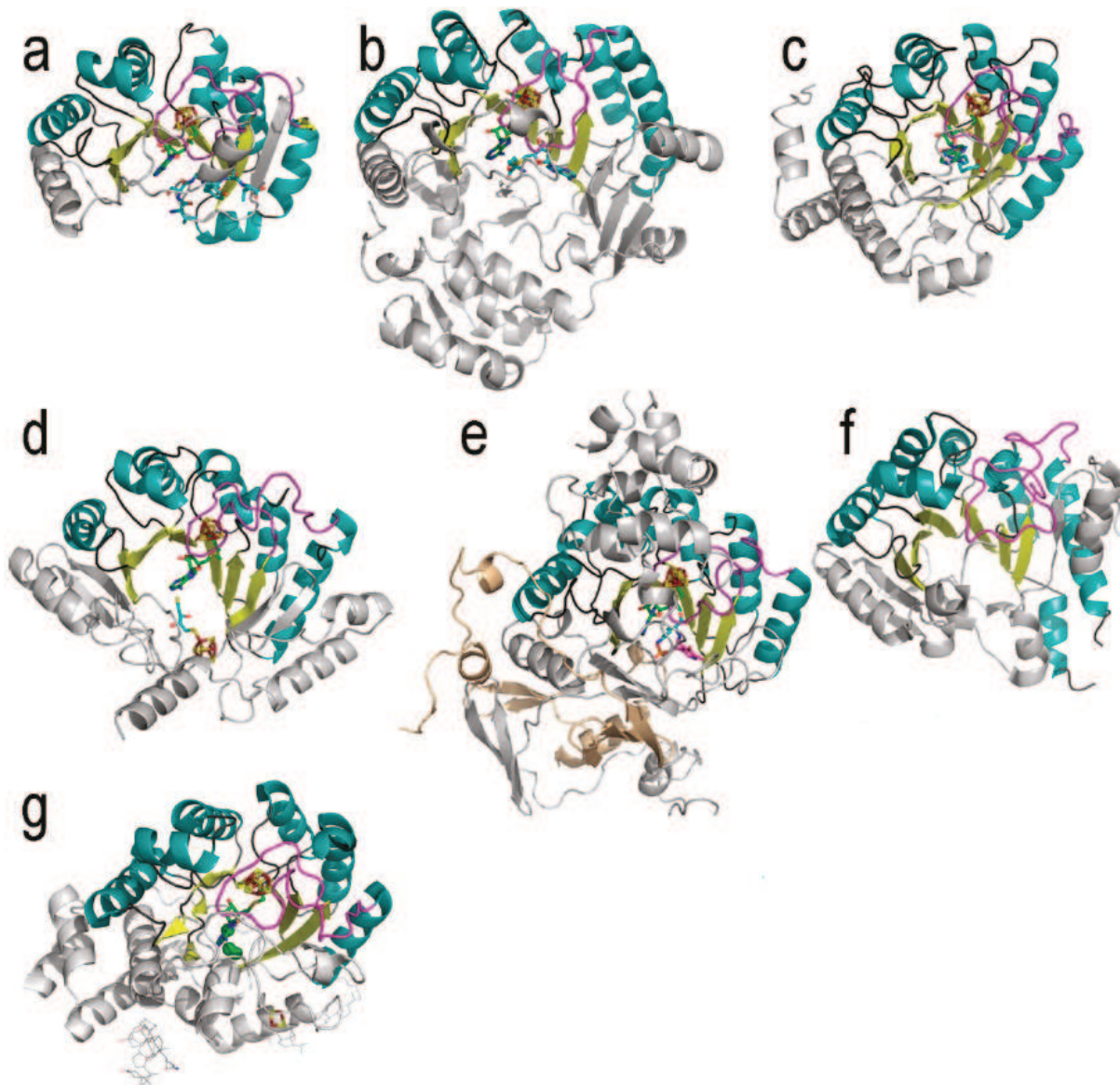
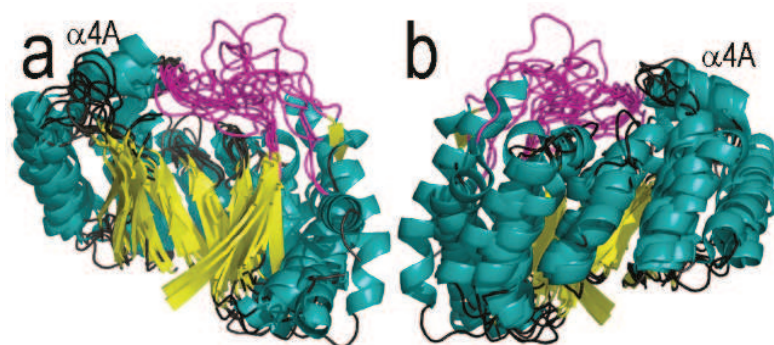


Figure III–2 Structures de protéines à radical SAM.

Figure tirée de Vey et Drennan (Vey and Drennan 2011). Vue du dessus de monomères de (a) PFL-AE, (b) HemN, (c) BioB, (d) MoaA, (e) LAM, (f) TYW1, (g) HydE. Le centre [4Fe-4S], la SAM, le substrat sont représentés en bâtons s'ils sont présents dans la structure. Le cœur ($\beta\alpha$)₆ est coloré de la façon suivante : hélices, bleu clair ; brins, jaune ; boucles, gris foncé ; boucle contenant le motif Cx₃Cx₂C liant le centre [4Fe-4S], rose. La SAM, le centre [4Fe-4S] et les substrats sont colorés de la façon suivante : fer, rouge foncé ; soufre, jaune ; carbones de la SAM, vert ; carbones de substrat, bleu clair ; oxygène, rouge ; azote, bleu. Les éléments de la protéine en dehors du cœur sont en gris clair. Dans (e), une partie de la seconde molécule de LAM est en couleur pêche et dans (g) trois atomes de chlore sont représentés par des sphères vertes.



des structures.

Figure III–3 Superposition de structures connues des PRS.

Figure tirée de Vey et Drennan (Vey and Drennan 2011). Superposition de la PFL-AE, HemN, BioB, MoaA, LAM, TYW1, HydE. Seul les cœurs sont représentés. Le code couleur est le même que dans la Figure III–2. (a) vue avant et (b) arrière

III.2) Des structures spéciales pour des PRS

Certaines PRS n'ont pas le type de structure décrit précédemment. Il s'agit des protéines impliquées dans la biosynthèse de la diphtamide (Dph2) (Zhang, Zhu et al. 2010) (Figure III–4 en haut), et dans la biosynthèse de la thiamine pyrimidine (ThiC) (Chatterjee, Li et al. 2008) (Figure III–4 en bas). La protéine Dph2 ne possède pas de structure en tonneau TIM mais une toute autre structure tertiaire. Quant à la protéine ThiC, elle possède un tonneau TIM mais il ne contient pas les 3 cystéines conservées fixant le centre [4Fe-4S].

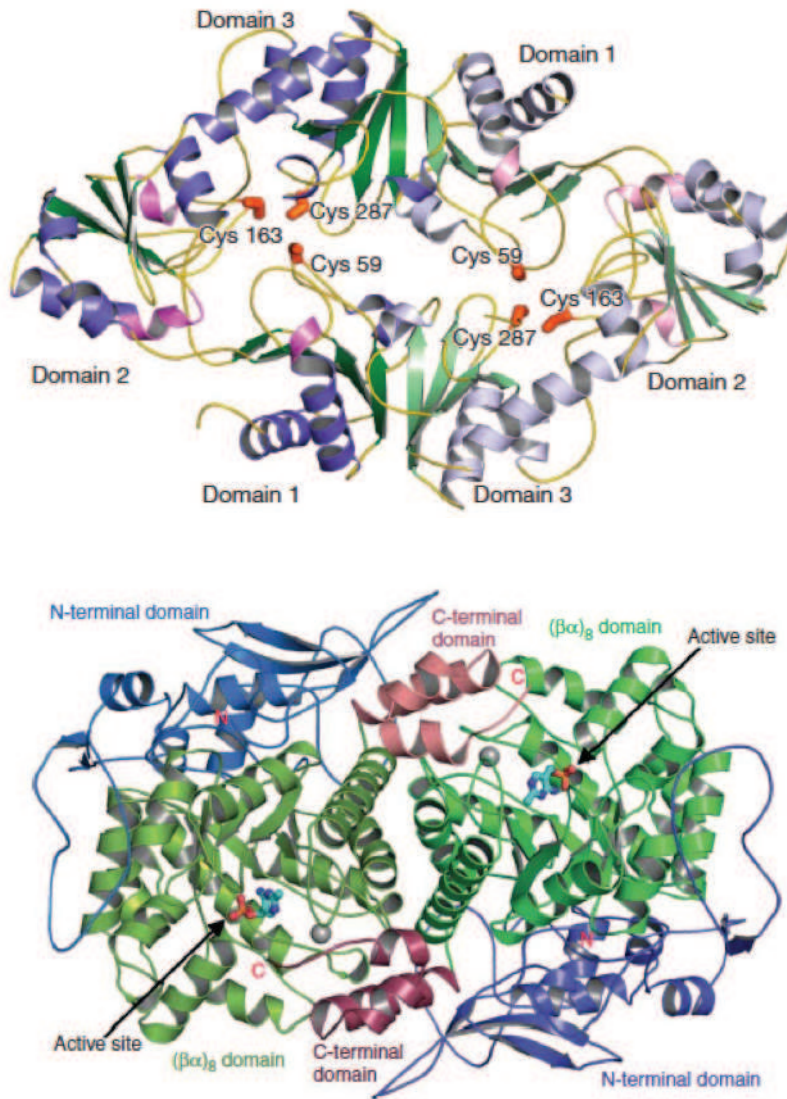


Figure III-4 Structure des homodimères Dph2 et ThiC.

Figures tirées de Zhang *et al.* (Zhang, Zhu et al. 2010) pour Dph2 (en haut) et (Chatterjee, Li et al. 2008) pour ThiC (en bas).

III.3) Le centre [4Fe-4S] et la liaison de la SAM

Le cœur (βα)₆ est proche du N-terminal des PRS. Le centre [4Fe-4S] est lié à une boucle émanant du haut du tonneau. A l'exception de la protéine ThiC (Figure III-4 en bas), cette boucle se trouve entre le brin β1 et l'hélice α1, et contient le motif Cx₃Cx₂C (Figure III-7). En ce qui concerne la protéine ThiC, ses 3 cystéines conservées se trouvent dans sa région C-terminal.

La SAM se lie à l'atome de fer libre du centre, par le carboxylate et l'amine de la méthionine. Ce centre, sous la forme [4Fe-4S]⁺, va fournir l'électron nécessaire au clivage de la SAM en radical 5'-désoxyadénosine (Ollagnier-de Choudens, Sanakis et al. 2002). Le

transfert d'électrons se ferait de l'atome de fer coordonnant la SAM vers l'atome de soufre du groupement méthionine de la SAM (Nicolet, Amara et al. 2009).

La SAM se trouve dans le haut du tonneau TIM partiel ou complet (Figure III–8). Les résidus qui la lient sont situés en majorité dans les brins β du cœur et les boucles qui les suivent. Ils sont donc très dispersés dans les séquences. L'amine de la méthionine de la SAM se lie près du C-terminal du brin $\beta 2$ et fait des liaisons hydrogène avec un motif structural « GGE » ou « riche en glycine » conservé (Nicolet and Drennan 2004) (Figure III–5 et Figure III–7). Dans le cas de BioB il s'agit des résidus A100, A101, W102 qui sont équivalents structurellement à ce motif (Figure III–5 et Figure III–7).

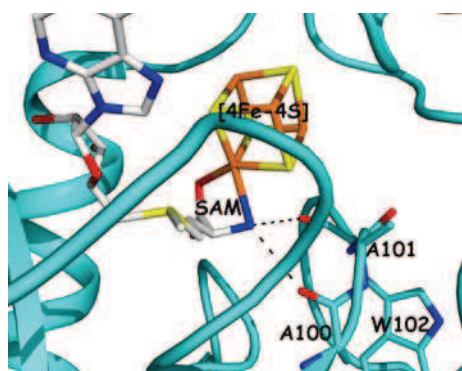


Figure III–5 Motif GGE.

De la protéine BioB (code PDB 1r30, (Berkovitch, Nicolet et al. 2004)).

Le groupement carboxylate de la méthionine est lié par des résidus de la protéine dont la localisation dans les séquences est variée. Les résidus liant le ribose de la SAM varient entre les différentes PRS, et la liaison à l'adénine se fait par un motif GxIxGxxE (Figure III–6 et Figure III–7). Dans le cas de BioB il s'agit des résidus G190, I192, G194 et E197.

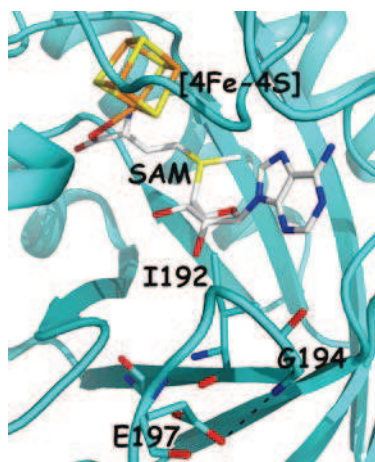


Figure III–6 Motif GxIxGxxE.

De la protéine BioB (code PDB 1r30, (Berkovitch, Nicolet et al. 2004)).

La position de la SAM par rapport au centre [4Fe-4S] est globalement similaire entre les différentes protéines.

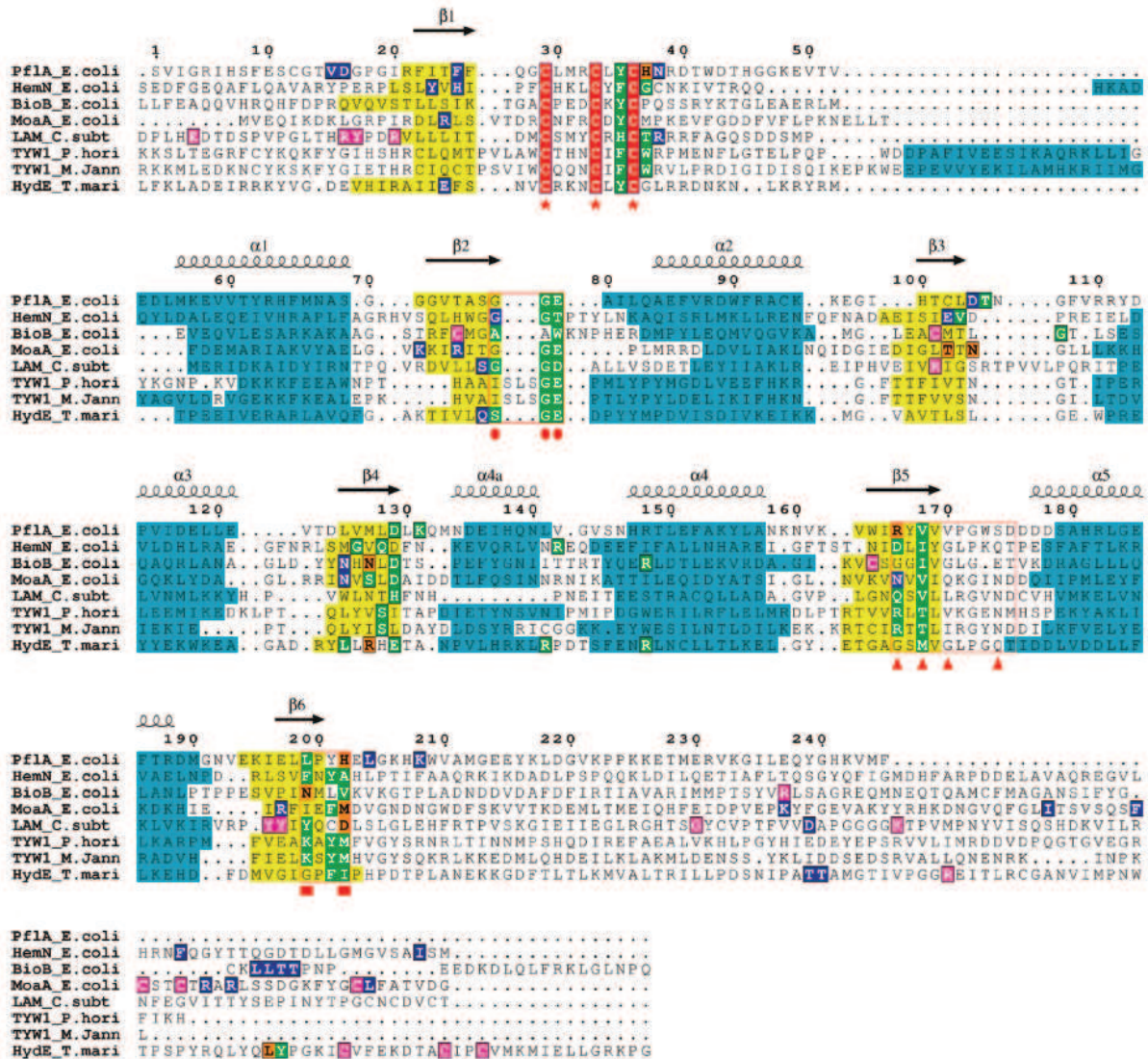


Figure III-7 Alignement structural de PRS.

Figure tirée de Vey et Drennan (Vey and Drennan 2011). Les séquences des protéines PFL-AE, HemN, BioB, MoaA, LAM, phTYW1, mjTYW1, et HydE dont une structure a été résolue, ont été alignées selon leurs éléments de structure secondaire du cœur radical SAM. Les principaux éléments de structure secondaire sont indiqués au-dessus des séquences. Le cœur radical SAM est indiqué par les couleurs bleu clair (hélices), jaune (brins). Les 3 cystéines conservées sont colorées en rouge. Les résidus en contact avec la SAM sont colorés en vert ; ceux en contact avec un substrat, en bleu ; ceux en contact avec un cofacteur, en rose ; ceux en contact avec la SAM et un substrat, en orange. Par clarté, certaines séquences n'ont été représentées que partiellement : HemN (33-331), BioB (23-313), MoaA (1-285), LAM (96-381), phTYW1 (30-342), mjTYW1 (30-311), et HydE (35-335). Les motifs conservés des PRS discutés sont représentés par : des étoiles, Cx₃Cx₂C ; cercles, GGE ; triangles, GxIxGxxE.

III.4) Fonction et réactivité des PRS

Le centre [4Fe-4S] est en général relié à la protéine par un motif hautement conservé Cx₃Cx₂C (Sofia, Chen et al. 2001). Il a été observé pour la protéine Dph2 (Zhang, Zhu et al. 2010) (Figure III–4 en haut), qu'elle lie un centre [4Fe-4S] par 3 cystéines de différents domaines (C59, C163 et C287) et qu'elle clive la SAM en radical 3-amino-3-carboxypropyl (ACP) et 5-méthylthioadénosine. Ceci pourrait être expliqué par sa structure très différente des autres PRS, mais un même clivage a été observé pour la glycérol déshydratase activase (Demick and Lanzilotta 2011), alors qu'elle contient un motif Cx₃Cx₂C. Les protéines Elp3 et ThiC (ses 3 cystéines se trouvent dans sa région C-terminal) possédant respectivement un motif Cx₄Cx₂C et Cx₂Cx₄C clivent la SAM en radical 5'-déoxyadénosine (5'-dA) et méthionine. La protéine HmdB contient un motif Cx₅Cx₂C. Elle produit de la 5'-dA (McGlynn, Boyd et al. 2010). La protéine aurait donc peut-être un autre moyen de spécifier sa réaction, qui est d'utiliser un motif différent de fixation du centre [4Fe-4S]. Il est donc intéressant de voir quel est l'effet de changement de motif sur la réaction d'initiation radicalaire.

La superposition des structures des différentes PRS montre que la SAM et le substrat se lient vers les régions C-terminal de chaque brin β du tonneau de la protéine (Figure III–8). La nature et la position des résidus qui lient le substrat sont responsables de la spécificité de l'enzyme et pourraient orienter la réaction radicalaire catalysée (Figure III–7). Il est intéressant de noter que la superposition des structures de différentes PRS contenant leur substrat et la SAM montre que la distance entre l'atome d'hydrogène qui sera arraché et l'atome C5' de la SAM où se trouvera le radical, est similaire (Vey and Drennan 2011).

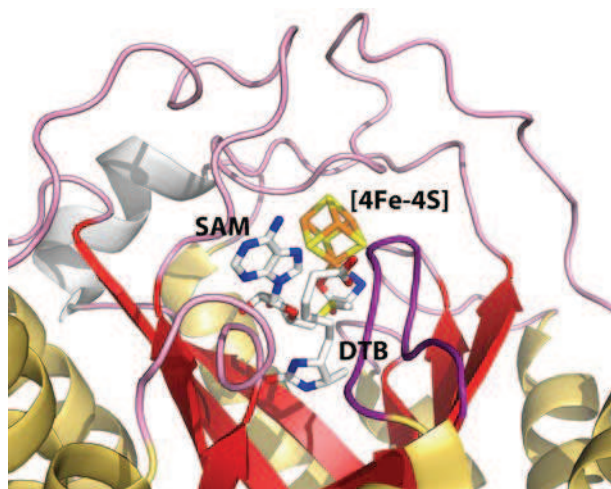


Figure III–8 Position du substrat et de la SAM dans les PRS.

Cas de la protéine BioB (code PDB 1r30, (Berkovitch, Nicolet et al. 2004)) avec son substrat la déthiobiotine (DTB).

III.5) Sous-famille de PRS ($\beta\alpha$)₈

Dans le cas des protéines HydE et BioB, deux motifs $\beta\alpha$ viennent compléter le cœur ($\beta\alpha$)₆, formant un tonneau triose phosphate isomérase (TIM) complet. Ce type de structure est adapté à lier des petits substrats (Nicolet and Drennan 2004), (Nicolet, Rubach et al. 2008). L'accès au site actif pourrait être contrôlé par une boucle « couvercle » qui couvre le tonneau et se trouve entre le brin β 8 et l'hélice α 8 ((Berkovitch, Nicolet et al. 2004), (Nicolet, Rubach et al. 2008)) (Figure III–9 et Figure III–10).

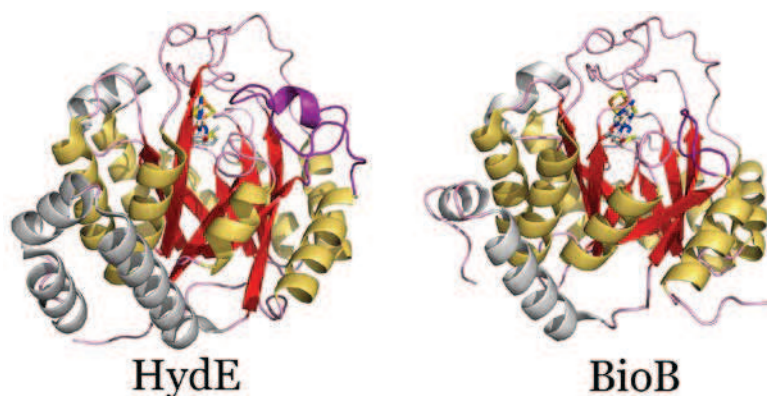


Figure III–9 Structures des PRS ($\beta\alpha$)₈ BioB et HydE.

BioB : code PDB 1r30, HydE : code PDB 3iiz. Les brins β et hélices α composant le tonneau triosephosphate isomérase sont représentés en rouge et jaune respectivement. Les hélices α supplémentaires sont en gris clair et les boucles en rose. La boucle fermant une extrémité du tonneau (« couvercle ») est représentée en

violet. Les centres [4Fe-4S] et la SAM sont représentés en bâtons.

La protéine BioB, biotine synthase, catalyse la réaction d'insertion de soufre dans la déthiobiotine (Frey, Hegeman et al. 2008) (Figure III–11). La protéine HydE fait partie de la machinerie de maturation de l'hydrogénase à fer-fer et serait impliquée dans la fabrication de la petite molécule organique pontante de son site actif (Mulder, Shepard et al. 2011). Les PRS ($\beta\alpha$)₈ ont des séquences similaires et possèdent une séquence conservée GxxP située en fin d'hélice α 8 (Nicolet, Martin et al. 2010) (Figure III–10). La protéine HmdB, supposée être impliquée dans la maturation d'Hmd (voir II.4)), possède aussi ce motif, et l'analyse des résidus G et P conservés qui seraient en début/fin de brin et hélice de sa séquence, comme il l'a été fait par Nicolet et Drennan (Nicolet and Drennan 2004), montre bien la présence de la structure du type TIM.

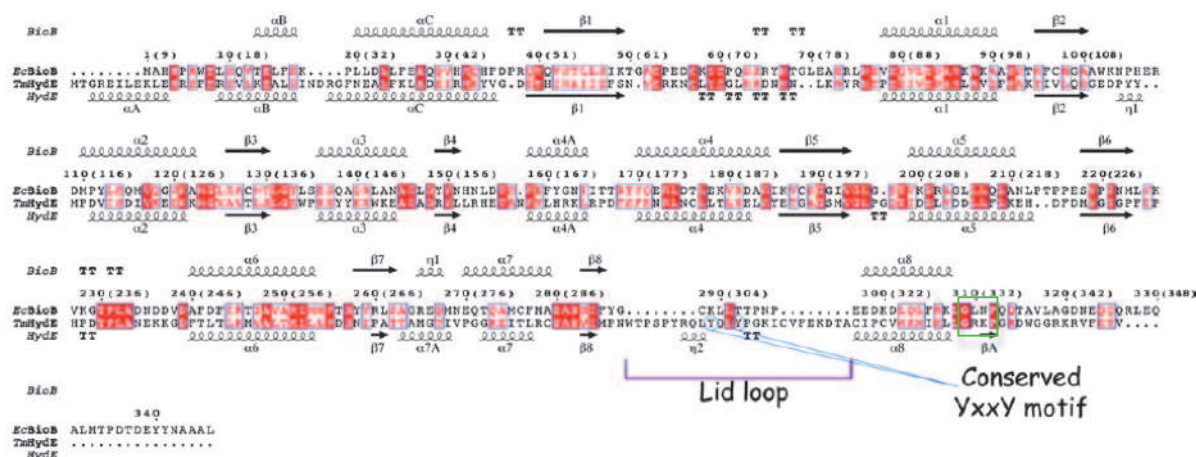


Figure III-10 Alignement structural des protéines HydE et BioB.

Figure tirée de Nicolet *et al.* (Nicolet, Rubach *et al.* 2008). Le couvercle est indiqué (lid loop). Le motif GxxP est encadré en vert.

Les PRS HydG, PylB, ThiH, CofG, CofH, NosL ont aussi ces caractéristiques de PRS ($\beta\alpha$)₈.

La protéine PylB (Gaston, Zhang *et al.* 2011) catalyse la réaction d'isomérisation de la lysine en transférant un groupement méthyle de la chaîne latérale de lysine pour produire le (3R)-3-méthyl-D-ornithine (Figure III-11). La structure de PylB isolée avec son produit de réaction a été très récemment résolue (Quitterer, List *et al.* 2011).

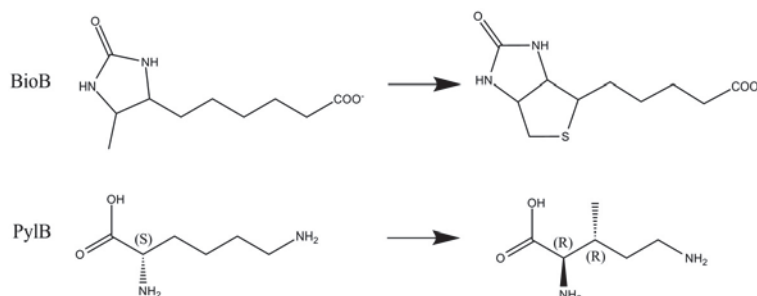
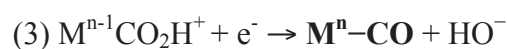
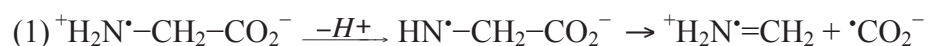


Figure III-11 Réactions catalysées par les protéines BioB et PylB.

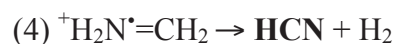
Les protéines ThiH, HydG et NosL catalysent la même réaction de clivage de la liaison C α -C β d'un résidu aromatique. La protéine ThiH, impliquée dans la biosynthèse de la thiamine pyrophosphate, est proposée produire du *p*-crésol et de la déhydroglycine ((Kriek, Martins *et al.* 2007) (Challand, Martins *et al.* 2010)) à partir de la tyrosine.

La protéine HydG, identifiée comme maturase de l'hydrogénase à [Fe-Fe] (Posewitz, King *et al.* 2004), clive la tyrosine pour former du *p*-crésol (Pilet, Nicolet *et al.* 2009). Les autres produits de la réaction sont le cyanure (Driesener, Challand *et al.* 2010) et le CO (Shepard,

Duffus et al. 2010), qui composent le site actif de l'hydrogénase à [Fe-Fe] (Figure I-6). La protéine HydG possède un domaine C-terminal contenant 3 cystéines conservées qui lient un deuxième centre [4Fe-4S]. La mutation de 2 des 3 cystéines en sérine a abolit la production de CO (Nicolet, Martin et al. 2010). Le centre [4Fe-4S] lié par ces cystéines est donc indispensable pour la production de CO. Cette même mutation des cystéines diminue la production de CN, et la mutation du domaine entier la diminue fortement. Néanmoins la production de *p*-cresol est elle-même diminuée dans le cas de la délétion de ce domaine (Tron, Cherrier et al. 2011). Le centre [4Fe-4S] n'est donc pas indispensable à la production de cyanure, mais on ne peut donc pas conclure sur le rôle de ce domaine dans sa production. La réaction de formation de cyanure ne nécessite donc pas de métal, contrairement à la production de CO. La formation simultanée du CO et CN nécessite de passer par un intermédiaire radical glycyll (Figure III-12). Les réactions proposées sont les suivantes (Nicolet, Martin et al. 2010):



avec le métal M^n qui serait là le deuxième centre [4Fe-4S]²⁺,



La protéine NosL (Zhang, Li et al. 2011), impliquée dans la biosynthèse du thiopeptide nocathiacine, clive le tryptophane en radical glycyll (observé) et 3-méthyl-indole (Zhang, Chen et al. 2011) pour ensuite produire l'acide 3-méthyl-2-indolique (Figure III-12).

Les protéines CofG et CofH composent la FO synthase (Graham, Xu et al. 2003). Cette enzyme transfère le groupement hydroxybenzyl du 4-hydroxyphenylpyruvate (HPP) sur le 5-amino-6-ribitylamino-2,4(1H,3H)-pyrimidinedione. Après oxydation, le produit est le 7,8-dideméthyl-8-hydroxy-5-deazariboflavin (FO). Une des deux protéines cliverait la liaison C_α-C_β d'un résidu aromatique comme les protéines ThiH, HydG et NosL. Le résidu aromatique est ici le HPP, un analogue de la tyrosine.

Un radical glycyll a été observé pour la protéine NosL (Zhang, Li et al. 2011), et un mécanisme proposé pour la production de CO et CN⁻ par la protéine HydG nécessite ce même intermédiaire : on peut donc proposer que ce même radical est l'intermédiaire de la réaction catalysée par la protéine ThiH pour ensuite former le thiazole avec la protéine ThiG. Dans le

cas de CofG ou CofH, le clivage de l'HPP produirait un radical *p*-cresol, avec le radical sur le groupement méthyl et du glyoxylate. Le radical *p*-cresol réagirait alors avec le 5-amino-6-ribitylamino-2,4(1H,3H)-pyrimidinedione (Figure III-12). Il serait donc intéressant d'avoir les structures de ces protéines/complexes de protéines afin d'élucider leurs mécanismes.

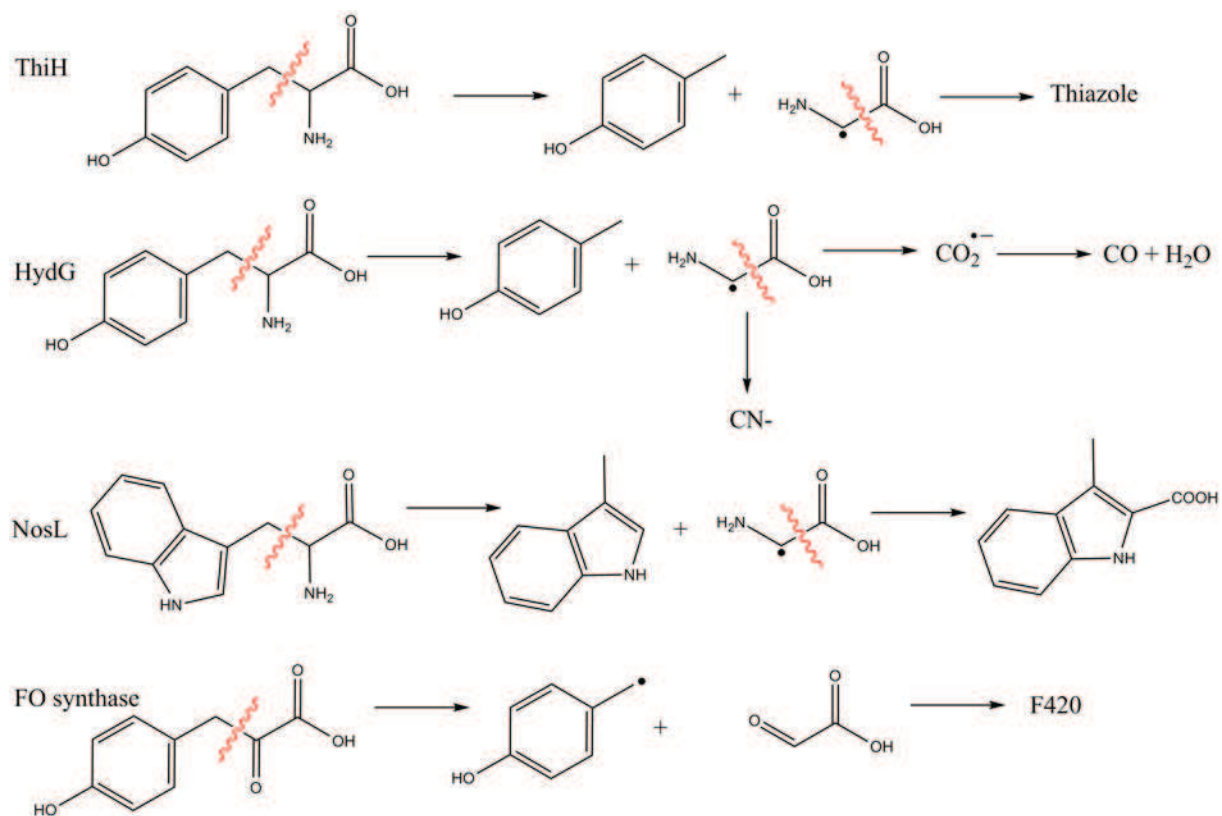


Figure III-12 Réactions catalysées par les protéines ThiH, HydG, NosL et la FO synthase.

Durant ma thèse j'ai étudié la maturation de sites métalliques de protéines. Je me suis intéressée d'une part, à la machinerie d'assemblage des centres fer-soufre (FeS) ISC et d'autre part aux protéines à radical *S*-Adénosyl-L-méthionine (PRS).

Pour le premier volet nous avons pu voir dans cette introduction qu'il existe de nombreuses protéines à centres FeS et que l'assemblage de ces centres ne se fait pas spontanément dans la cellule, que ce soit chez les eucaryotes ou les bactéries. Il existe des machineries protéiques pour fabriquer les centres FeS et les transférer aux apoprotéines cibles. Chez les bactéries, la machinerie ISC est la machinerie générale d'assemblage des centres FeS. C'est sur la protéine échafaudage IscU que se fabrique le centre FeS, le fer étant apporté par une protéine donneuse de fer et le soufre par la cystéine désulfurase IscS. Un centre [2Fe-2S] est assemblé sur IscU. Pour étudier le mécanisme d'assemblage de ce centre j'ai cherché à déterminer la structure par cristallographie des rayons X d'un complexe IscS-IscU obtenu *in vivo* et contenant un centre [2Fe-2S].

Pour le second volet nous avons pu voir que les PRS constituent une superfamille catalysant des réactions très variées. Certaines d'entre elles ont un rôle de maturation de sites actifs de protéines, telles que les protéines HydE et HydG pour la machinerie de maturation de l'hydrogénase à fer-fer et NifB pour la machinerie de maturation de la nitrogénase. Les gènes des protéines participant à l'assemblage de ces sites se situent en général à proximité du gène codant pour l'apoprotéine.

Dans le cas de l'hydrogénase à fer Hmd, impliquée dans le processus de méthanogenèse, un gène codant pour la protéine HmdB se trouve à proximité de son gène codant, et uniquement dans les organismes possédant Hmd. Nous avons donc émis l'hypothèse que la protéine HmdB serait impliquée dans la biosynthèse du cofacteur FeGP de Hmd. Peu de travaux ont été réalisés sur le mécanisme d'assemblage du site actif d'Hmd. Néanmoins nous avons supposé qu'il existe aussi une machinerie pour la biosynthèse et l'insertion de son cofacteur, d'autant plus que ce dernier peut être inséré *in vitro* dans une apoprotéine. Pour étudier la maturation de la protéine Hmd, j'ai donc réalisé durant ma thèse une étude structurale de la protéine HmdB.

Matériel et méthodes

I) Biologie moléculaire

I.1) Clonage dans le plasmide pETDuet-1 : plasmides pETMK0016Duet-1 et pETMJ0785Duet-1

I.1.1) Amplification des gènes d'intérêt

La protéine HmdB de *Methanopyrus kandleri* AV19 est codée par le gène MK0016 (locus primaire) de 1071 pb (locus JCVI : NT01MK0018, code d'accès Uniprot Q8TZB9, identifiant GenBank AAM01233.1). Son gène a été inséré préalablement dans un plasmide pET28a (pMK0016), disponible au laboratoire.

HmdB de *Methanococcus jannaschii* DSM2661 est codée par le gène MJ0785 de 1128 pb (locus JCVI : MJ_0785, code d'accès Uniprot Q58195, identifiant GenBank AAB98782.1) et partage 55% d'identité de séquence avec son orthologue de *Methanopyrus kandleri*. Le gène a lui aussi déjà été inséré dans le plasmide pET28a (pMJ0785), disponible au laboratoire.

Afin de co-exprimer ces protéines avec d'autres protéines dont les gènes sont insérés dans un plasmide pRSFDuet-1, les gènes des protéines HmdB sont réinsérés dans le site MCS2 du plasmide pETDuet-1 (Novagen) (5420 pb), qui ne contient pas d'étiquette (Tag), entre les sites NdeI et KpnI pour MK0016 et entre les sites NdeI et XhoI pour MJ0785 (Figure I-1).

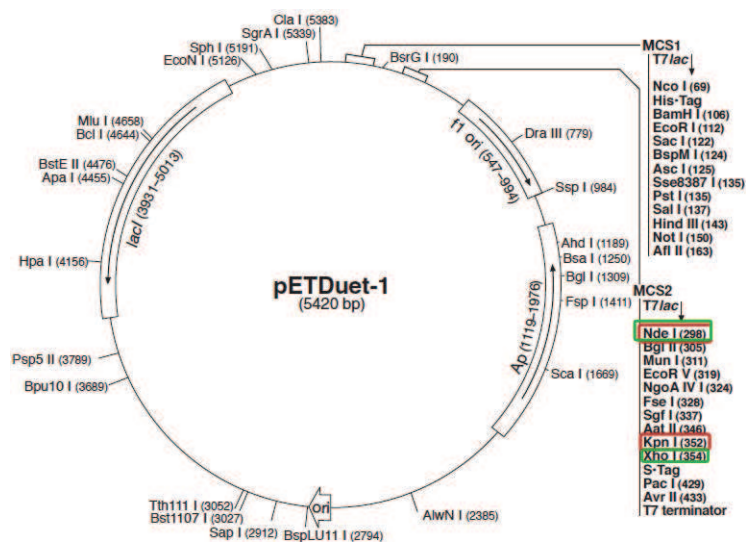


Figure I-1 Insertion des gènes MK0016 et MJ0785 dans le plasmide pETDuet-1.

Sites de restriction d'insertion de MJ0785 en vert et MK0016 en rouge.

Les gènes des protéines HmdB sont amplifiés par PCR. Les séquences des amorces utilisées pour amplifier les gènes codant pour HmdB dans les plasmides pET28a sont données dans le Tableau I-1. Elles contiennent à leurs extrémités les sites de coupure des enzymes de restriction qui correspondent aux sites d'insertion dans le pETDuet-1.

gène	nom	séquence des amorces	Tm (°C)
MK0016	mk0016_for ^(a)	5'GGTAATCCATATGCGTTTCAAAGACGCGCT ^{3'} (NdeI)	65,9
	mk0016KpnI_rev ^(b)	5'GGTAATCCGGTACCTTATCAGCTCGAGCGCTGCGT ^{3'} (KpnI)	67,2
MJ0785	mj0785_for ^(a)	5'GGTAATCCATATGAAATATACTTATAAAATAAGTAAAGT AAAAATATCTGGTG ^{3'} (NdeI)	62,3
	mj0785XhoI_rev ^(b)	5'GGTAATCCTCGAGTTATTATTTCCATAAATCTTCAAATA ATTCTCATACC ^{3'} (XhoI)	63,5

Tableau I–1 Amorces utilisées pour insérer HmdB dans le pETDuet-1.

^(a) amorce sens. ^(b) amorce antisens. Les sites de coupure des enzymes de restriction (entre parenthèses à côté de la séquence) sont indiqués en gras dans la séquence.

Les mélanges réactionnels pour la PCR et programmes PCR sont donnés dans le Tableau I–2.

Mélange réactionnel			
volume final de 50 µL (complété à l'eau stérile)			
amorce sens (voir Tableau I–1)		15 pmol	
amorce antisens (voir Tableau I–1)		15 pmol	
tampon 5x Phusion™ GC (Finnzymes), (version HF pour MJ0785)		10 µL	
dNTP (à 10 mM)		1 µL	
pMK0016 ou pMJ0785		120 ng /160 ng	
Phusion™ DNA Polymerase (Finnzymes)		1 U	
Programme PCR			
dénaturation initiale		98°C	5 min
dénaturation	35 cycles	98°C	30 s
hybridation		64°C (62°C pour MJ0785)	30 s
élongation		72°C	45 s
élongation finale		72°C	10 min

Tableau I–2 Amplification des gènes par PCR. Mélange réactionnel et programme.

I.1.2) Digestion enzymatique de l'insert et du plasmide

20 µL du produit PCR (insert) sont digérés pendant 2 h à 37°C dans un volume final de 60 µL complété à l'eau stérile avec les enzymes de restriction correspondantes, 6 µL de tampon associé et 1 µL de BSA (NEB, à 10 mg/mL). En parallèle 1500 ng de plasmide pETDuet-1 sont digérés dans les mêmes conditions.

Les enzymes utilisées pour MK0016 sont 30 U de NdeI (New England BioLabs, NEB) et 15 U de KpnI (NEB) ; le tampon utilisé est le 10x NEBuffer 1, et pour MJ0785 : NdeI et XhoI (30 U, NEB) et le tampon NEBuffer 4.

I.1.3) Traitement à la phosphatase

L'Antarctic Phosphatase catalyse l'élimination des groupements 5' phosphate d'ADN et ARN. Les fragments ainsi traités ne peuvent pas s'auto-liaison lors de la présence d'une ligase. Avant l'étape suivante de ligation, le plasmide digéré est traité avec 5 U d'Antarctic Phosphatase (NEB), 8 µL de tampon 10x Antarctic Phosphatase (le volume est complété à 80 µL en H₂O stérile) pour une durée de réaction de 30 min, à 37°C.

Les deux tubes réactionnels sont chauffés à 65°C pendant 20 min afin d'éliminer les enzymes. L'insert et le plasmide digérés sont purifiés sur colonne Montage® (Millipore) (en ajoutant 200 µL d'eau au produit PCR).

I.1.4) Ligation de l'insert et du plasmide digérés

La concentration des inserts et plasmides est déterminée en mesurant l'absorbance à 260 nm (cuve en quartz) puis en utilisant la formule $c \text{ (ng/}\mu\text{L)} = A_{260} \cdot 50 \cdot (1/\text{dilution})$. Le plasmide et l'insert (volume total de 10 µL avec un excès insert/vecteur de 20, dilutions au préalable à faire si nécessaire) sont ajoutés à 10 µL de tampon T4 DNA ligation 2x et 1 µL de T4 DNA Ligase (Rapid DNA Ligation Kit, Roche) pour une incubation d'une heure à température ambiante.

I.1.5) Transformation du produit de ligation

Le produit de ligation (20 µL) est ajouté à 100 µL de cellules DH5α (préalablement traitées pour les rendre « compétentes ») pour être transformées. Les cellules en présence du plasmide sont incubées 30 min sur glace puis 1 min à 42°C pour le choc de température et à

nouveau sur glace pendant 2 min. 250 μ L de milieu de culture LB sont ensuite ajoutés (Sigma) et le mélange est incubé 1h à 37°C avec une agitation de 200 rpm. 250 μ L de liquide sont éliminés après centrifugation et le restant est resuspendu et étalé sur boîte LB-Agar (Sigma) contenant de l'ampicilline à 50 μ g/mL. La boîte est mise à l'incubateur à 37°C pendant une nuit.

I.1.6) Test PCR sur colonies et extraction de plasmide

Un test PCR sur colonies est réalisé afin d'identifier celles possédant l'insert. Pour cela on pique différentes colonies (environ 10) que l'on resuspend dans 10 μ L d'eau stérile. A 2 μ L de ces solutions on ajoute 2 μ L de tampon 10x Taq (NEB), 0.5 μ L de dNTP (10 mM), 10 pmol d'amorces sens et anti-sens (voir Tableau I-1) et 2.5 U de Taq DNA polymérase (NEB). On complète le volume à 20 μ L avec de l'eau stérile. Les tubes sont chauffés 2 min à 95°C, puis subissent 30 cycles de 30 s à 95°C / 30 s à 64°C (62°C pour MJ0785) / 1 min 30 s à 72°C et en dernière étape 10 min à 72°C. Durant le temps de la PCR les 8 μ L restants de colonie piquée sont resuspendus dans 1 mL de LB et mis à 37°C avec une agitation de 200 rpm. On réalise un gel d'agarose 1% sur les produits PCR. Un clone est dit positif si l'on retrouve la taille attendue de l'insert. A quelques cultures de clones positifs on ajoute 5 mL de LB contenant de l'ampicilline à 50 μ g/mL qu'on incube à 37°C, 200 rpm pour une nuit. Le lendemain on réalise une extraction de plasmide après avoir centrifugé les tubes à 5000 rpm pendant 15 min à 4°C (Nucleospin® Plasmid, Macherey-Nagel) en réalisant l'étape optionnelle de lavage avec le tampon AW et en chauffant le tampon AE à 70°C avant de l'utiliser. Les échantillons sont séquencés afin de vérifier leur justesse.

Les plasmides pETMK0016Duet-1 et pETMJ0785Duet-1 sont obtenus.

I.2) Mutant C-terminal

Un mutant de HmdB remplaçant la Gly336 par un codon stop est réalisé par mutagenèse dirigée sur le plasmide pETMK0016Duet-1 grâce aux oligonucléotides suivant :

nom	séquence des amorces	Tm
Hmdb_cut_for ^(a)	5' CCGAGCTCAGGT G AGACCGAGCTTGC 3'	67°C
Hmdb_cut_rev ^(b)	5' GCAAGCTCGGTCTCACCTGAGCTCGGG 3'	67°C

Tableau I-3 Amorces utilisées pour muter HmdB au codon 336.

^(a) amorce sens. ^(b) amorce antisens. Le site de mutation est indiqué en gras dans la séquence des amorces.

La mutagenèse est faite sur 120 ng de pETMK0016Duet-1 en solution avec 15 pmol d'amorces Hmdb_cut_for et Hmdb_cut_rev (décrites dans le Tableau I-3), 1 µL de dNTP, 10 µL de tampon 5x PhusionTM HF et 1 U de PhusionTM DNA Polymerase (Finnzymes). Le volume final est de 50 µL complété à l'eau stérile. L'échantillon est alors chauffé 5 min à 98°C, puis subit 25 cycles de 30 s à 98°C / 30 s à 67°C / 1 min 30 s à 72°C pour finir avec une dernière étape de 10 min à 72°C.

Les ADN méthyltransférases se trouvent dans une grande variété de procaryotes et eucaryotes, elles transfèrent un groupement méthyl de la S-adénosylméthionine à un résidu adénine ou cytosine lors de la réplication de l'ADN. L'ADN provenant d'*E. coli* (transformations de cellules DH5α) est méthylé, alors que l'ADN produit par PCR ne l'est pas. L'enzyme DpnI clive les séquences GATC méthylées (GA(Me)TC), elle peut donc être utilisée pour éliminer l'ADN non produit par PCR. Le produit PCR est digéré pendant 1h30 à 37°C par 30 U de DpnI dans 6 µL de tampon 10x NEBuffer 4 (NEB), complété à 60 µL à l'eau stérile. Un contrôle est réalisé en parallèle afin de vérifier l'efficacité de l'enzyme, en traitant 120 ng de pETMK0016Duet-1 avec 30 U de DpnI dans 6 µL de tampon 10x NEBuffer 4, complété à 60 µL à l'eau stérile. Des cellules DH5α sont transformées par 10 µL des deux produits (voir le protocole de transformation dans le paragraphe I.1.5)). Les produits de transformation sont étalés sur boîte LB/Agar contenant de l'Amp à 50 µg/mL. Après une nuit d'incubation à 37°C, quelques colonies isolées sont inoculées dans 5 mL de LB contenant de l'Amp à 50 µg/mL et incubées à 37°C, 200 rpm pour une nuit. Le lendemain une extraction de plasmide est réalisée après avoir centrifugé les tubes à 5000 rpm pendant 15 min à 4°C (Nucleospin® Plasmid, Macherey-Nagel), en réalisant l'étape optionnelle de lavage avec le tampon AW, et en chauffant le tampon AE à 70°C avant de l'utiliser. Les échantillons sont alors séquencés afin de vérifier leur justesse.

Le plasmide pETMK0016_cut_Duet-1 est obtenu.

I.3) Co-transformation de pETMK0016Duet-1, pETMK0016_cut_Duet-1, pETMJ0785Duet-1 avec pRSFIscDuet-1 ou pRSFIscMetKDuet-1

Comme il a été discuté dans l'introduction, les centres FeS bactériens sont synthétisés par la machinerie ISC codée par un opéron comportant les protéines IscR, IscS, IscU, IscA, HscB, HscA, fdx et IscX. La protéine IscR est régulatrice de l'opéron. Lorsque l'on surexprime une protéine bactérienne contenant un centre FeS il est donc intéressant de le faire avec la

machinerie ISC en parallèle. De cette façon la protéine surexprimée devrait déjà posséder son centre FeS, ceci permettrait d'éviter sa reconstitution *in vitro*. Le plasmide pRSFIscDuet-1 (Figure I-2) est un plasmide pRSFDuet-1 (Novagen) dont l'opéron codant pour la machinerie Isc d'*E. coli* K12-MG1655 a été inséré, sans le régulateur IscR (b2530 à b2524, 5020 pb) dans le site de clonage MCS1, entre les sites d'enzymes de restriction NcoI et NotI. Dans les protéines à radical SAM, le cofacteur se lie au centre [4Fe-4S] au niveau de l'atome de fer unique, ce qui le stabilise. La SAM est synthétisée *in vivo* par l'enzyme SAM synthétase (MetK) à partir de L-méthionine, d'eau et d'ATP. Il est donc intéressant de surexprimer MetK en plus de la machinerie ISC sans IscR. Le plasmide pRSFIscMetKDuet-1 (Figure I-2) possède en plus le gène codant pour la SAM synthétase d'*E. coli* K12-MG1655 (b2942). Il a été inséré (1155 pb) dans le site de clonage MCS2 entre les sites d'enzymes de restriction NdeI et XhoI. Il a été obtenu au laboratoire dans un second temps ce qui explique pour nous ne l'avons pas utilisé pour toutes les co-transformations.

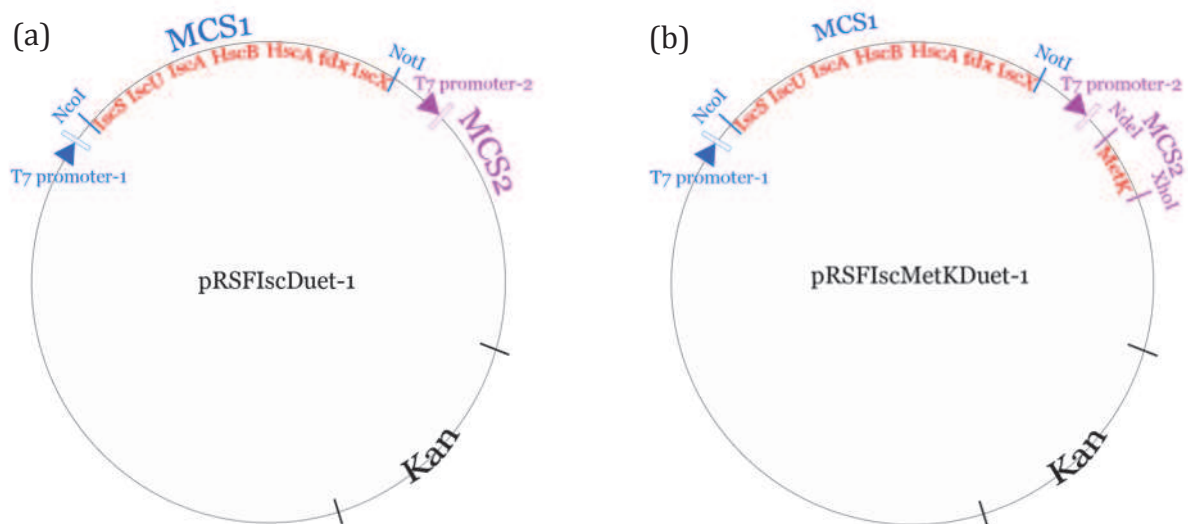


Figure I-2 Représentation schématique des plasmides pRSFIscDuet-1 et pRSFIscMetKDuet-1.

(a) plasmide pRSFIscDuet-1, (b) plasmide pRSFIscMetKDuet-1.

Les bactéries ont été co-transformées par les plasmides pRSFIscDuet-1 et pRSFIscMetKDuet-1, et les plasmides obtenus précédemment :

- 100 μ L de cellules BL21 (DE3) par 360 ng de pETMK0016Duet-1 et 500 ng de pRSFIscDuet-1 (voir le protocole de transformation dans le paragraphe I.1.5) et le produit de transformation est étalé sur boîte LB-Agar (Sigma) contenant de l'Amp

à 25 µg/mL et de la Kan à 15 µg/mL. La boîte est mise à l'incubateur à 37°C pendant une nuit (HmdB native de *Methanopyrus kandleri*).

- des cellules BL21 (DE3) sont co-transformées par 130 ng de pETMK0016_cut_Duet-1 et pRSFiscMetKDuet-1 selon le même protocole que ci-dessus (HmdB mutée de *Methanopyrus kandleri*).
- des cellules BL21 (DE3) contenant le plasmide pRSFiscMetKDuet-1 sont transformées par 180 ng de pETMJ0785Duet-1 (voir le protocole de transformation dans le paragraphe I.1.5)), et le produit est étalé sur boîte LB-Agar (Sigma) contenant de l'Amp à 50 µg/mL et de la Kan à 30 µg/mL. La boîte est mise à l'incubateur à 37°C pendant une nuit (HmdB de *Methanococcus jannaschii*).

II) Biochimie

II.1) Protéine HmdB

II.1.1) Surexpression

On réalise une préculture de 8 h à 37°C, avec une agitation de 200 rpm, des colonies obtenues lors de la transformation des cellules BL21 (DE3) (voir le paragraphe I.3)) en piquant plusieurs colonies que l'on resuspend dans 5 mL de milieu LB, Kan 30 µg/mL, Amp 50 µg/mL (deux précultures sont réalisées au cas où l'une des deux ne croîtrait pas). 5 mL de préculture sont dilués dans 200 mL de milieu LB dans lequel on ajoute à nouveau de la Kan et de l'Amp pour être à une concentration de 30 µg/mL et 50 µg/mL respectivement. On incube une nuit à 37°C, avec une agitation de 200 rpm. Le lendemain 20 mL de préculture sont prélevés et dilués dans 1 L de milieu TB (Sigma) (réalisé 6 fois, 6 * 1 L de culture au total), en maintenant les mêmes concentrations en antibiotiques. On incube à 37°C, 200 rpm jusqu'à obtenir une DO à 600 nm de 0.6 à 0.8. Lorsque cette DO est atteinte on induit à l'aide d' 1 mM d'IPTG en supplémentant le milieu avec 10 µM de cystéine, 0.1 g/L de citrate de fer et on incube 3 h à 37°C, 200 rpm. On centrifuge à 6000 rpm pendant 15 min à 4°C pour récupérer les culots que l'on transfère dans 3 tubes Falcon (3 * 2L de culture) (centrifugation à 4100 rpm pendant 15 min à 4°C et élimination du surnageant) pour congélation à - 80°C.

II.1.2) Purification de la protéine HmdB de *Methanopyrus kandleri*

La protéine HmdB de *Methanopyrus kandleri* native ou mutée se purifie d'après le protocole suivant : les culots cellulaires obtenus après culture sont resuspendus dans 40 mL de 50 mL de (Tris 20 mM pH 7.5, 10 mM DTT, 1mM MgCl₂, 1 tablette cOmplete EDTA-free (Roche)). Les bactéries sont lysées par sonication pendant 2 min à 90 % d'amplitude (2 s de sonication, 10 s de pause). On ajoute 500 U de Benzonase® nuclease (Novagen) et on incube 30 min à température ambiante.

La première étape de purification se fait par un chauffage au bain-marie à 70°C pendant 20 min. On centrifuge à 15000 rpm pendant 35 min afin de séparer les débris cellulaires de la protéine se trouvant dans la partie soluble. Toutes les étapes précédentes ont été faites à l'air. Les tubes de centrifugation sont rentrés en BAG. Les étapes suivantes sont alors toutes été faites en BAG.

Le surnageant est injecté sur une colonne d'affinité HiTrap™ Blue HP V = 5 mL (résine Blue Sepharose, GE Healthcare). La molécule greffée sur la résine est le bleu Cibacron F3G-A, qui lie les protéines à radical SAM HydE et BioB (alors que la protéine HydG, une autre protéine à radical SAM ne se lie pas). Etant donné la similarité de séquences de la protéine HmdB avec HydE et BioB une affinité pour cette colonne a été attendue, ce qui est le cas.

Le tampon d'équilibration (A) est du Tris 20 mM pH 7.5. L'échantillon est chargé avec un débit de 0.5 mL/min et la colonne est ensuite lavée avec 7 CV de tampon (A). Un gradient de 0 à 100 % en tampon (A) + 1 M NaCl en 10 CV suit ce lavage. Le lavage et l'élution se font à débit de 1 mL/min. Les fractions contenant la protéine sont identifiées sur le chromatogramme et par SDS-PAGE, et sont rassemblées pour être concentrées sur cellule Amicon membrane 10 K (Millipore) jusqu'à une concentration d'environ 5 mg/mL. La protéine est pure à ce stade.

Un dosage au rose bengal est fait afin d'évaluer la concentration en protéine, suivi d'un dosage en fer pour évaluer le contenu en fer de la protéine. Si ce contenu n'est pas suffisant, ce qui a été le cas, les centres FeS de la protéine sont reconstitués par ajout de 5 mM de DTT, agitation pendant 30 min, ajout de Na₂S 6 fois en excès, agitation pendant 30 min et ajout de FeCl₃ 6 fois en excès (ou (NH₄)₂Fe(SO₄)2.6H₂O). Un premier essai de reconstitution avec des excès en Na₂S et source de Fe de 10 montrait un résultat similaire ; leur quantité a alors été diminuée afin d'éviter un excès de précipités. Les poudres des composés ajoutés à la protéine ont été dissoutes en BAG juste avant utilisation. La solution est laissée 1 h 30 sous agitation, centrifugée pour éliminer les précipités, et 5 équivalents de SAH sont ajoutés au surnageant

protéique. La SAH est ajoutée afin de stabiliser le centre FeS formé ; la SAH est d'ailleurs un inhibiteur des PRS. Après 20 min d'incubation la solution est injectée sur une colonne HiTrap Desalting V = 5 mL (GE Healthcare), équilibrée avec du Tris 20 mM pH 7.5 300 mM NaCl avec un débit de 2 mL/min. La protéine est concentrée sur cellule Amicon membrane 30 K. Un dosage au rose bengal est fait afin d'évaluer la concentration en protéine, suivi d'un dosage en fer pour évaluer le contenu en fer de la protéine. La protéine est divisée en aliquots et congelée dans l'azote liquide.

II.1.3) Essai de purification de la protéine HmdB de Methanococcus jannaschii

Des essais de purification de la protéine HmdB de *Methanococcus jannaschii* ont été faits. Les culots cellulaires obtenus après culture ont été resuspendus dans 40 mL de 50 mL de (Tris 20 mM pH 7.5, 10 mM DTT, 1mM MgCl₂, 1 tablette cOmplete EDTA-free (Roche). Les bactéries sont lysées par sonication pendant 2 min à 90 % d'amplitude (2 s de sonication, 10 s de pause). 500 U de Benzonase® nuclease (Novagen) ont été ajoutés, puis l'échantillon a été incubé 30 min à température ambiante.

La première étape de purification a été faite (ou non) par un chauffage au bain-marie à 70°C pendant 20 min. On a centrifugé à 15000 rpm pendant 35 min afin de séparer les débris cellulaires de la protéine, laquelle se trouve dans la fraction soluble. Toutes les étapes précédentes ont été faites à l'air. Les tubes de centrifugation sont rentrés en BAG. Les étapes suivantes sont alors toutes été faites en BAG.

Le point isoélectrique de la protéine est de 8.65 (contre 5.72 pour la protéine de *Methanopyrus kandleri*). La deuxième étape de purification choisie est une colonne échangeuse de cations avec la résine SP Sepharose Fast Flow (réf. 170729-10, GE Healthcare, V = 25 mL (section de 16 mm)), à un pH de 6.5. La plupart des protéines étant acides (pI inférieur à 7), elles ne se lieront pas à la résine, puisqu'elles seront chargées négativement à ce pH. Le tampon d'équilibration (A) est du BisTris 50 mM pH 6.5. Le débit utilisé est de 2 mL/min, et, après la charge, la colonne a été lavée avec 3 CV de tampon (A). Un gradient de 0 à 100 % en tampon (A) + 1 M NaCl en 6 CV a suivi ce lavage.

II.2) Complexe IscSU

II.2.1) Production et purification de la protéine AflScS et du complexe AflScSU

La protéine IscS (locus du gène : AF0564) d'*Archaeoglobus fulgidus* a un poids moléculaire de 41.8 kDa, un pI de 5.7. Elle comporte 382 acides aminés.

La protéine IscU du même organisme (locus du gène : AF0565) a un poids moléculaire de 16.9 kDa, comporte 153 acides aminés. Son pI est de 5.46.

Le gène AF0564 ou le segment d'ADN contenant les gènes AF0564 et AF0565 ont été clonés dans un plasmide d'expression pAra13, sans ajout d'étiquette au C-term ou N-term des protéines. Une mutagenèse dirigée a alors été faite sur le plasmide contenant les gènes AF0564 et AF0565 afin d'obtenir un mutant Asp35Ala d'IscU (AF0565-D35A).

Les cellules d' *E. coli* C41 (DE3) ont été transformées par un des deux plasmides contenant le gène AF0564, ou AF0564 et AF0565-D35A. La purification de IscS et du complexe (IscS-D35A-IscU)₂ a été faite en boîte à gants en condition anaérobie au laboratoire du Pr. Dennis Dean (Virginia Tech, Etats-Unis) collaborateur du projet.

Nous avons obtenu les structures du complexe (IscS-D35A-IscU)₂ traité au DTT (DT), exposé à l'air (AE) et non traité (UT).

II.2.2) Calculs théoriques

Des calculs théoriques concernant le complexe IscSU ont été faits avec la suite de programmes Schrödinger (Schrödinger 2011), par le Dr Patricia Amara de notre équipe.

La modélisation d'une métalloprotéine nécessite d'utiliser la mécanique quantique pour décrire les changements de la structure électronique du site métallique considéré. Toutefois, les calculs de mécanique quantique sont coûteux en temps de machine, ce qui limite considérablement la taille des systèmes qu'il est possible d'examiner. Un moyen pour étudier efficacement les réactions dans les protéines est de former des potentiels hybrides couplant la mécanique quantique (MQ) et la mécanique moléculaire (MM) pour décrire le système protéique ((Quitterer, List et al. 2011), (Hendrickson and Leigh 2008)). La mécanique moléculaire considère la molécule comme un ensemble de particules classiques. Les atomes d'un système sont définis comme des sphères reliées entre elles par des ressorts, et l'énergie totale est écrite de manière empirique grâce à un ensemble de paramètres constituant un

champ de force. Le site actif ou site d'intérêt est modélisé par la mécanique quantique où les atomes sont décrits comme des noyaux et des électrons, alors que le reste de la matrice protéique est décrit par la mécanique moléculaire.

Afin de voir s'il y a assez de place pour un acide aspartique à la place de l'Ala35 dans notre structure du complexe DT *Af* (IscS-Ala₃₅-IscU)₂, un modèle dérivé de cette structure a été construit. Les atomes d'hydrogène ont été construits. Un composant IscS-IscU a été extrait et l'Ala35 a été substituée par un Asp, l'orientation de sa chaîne latérale a été optimisée, ainsi que la boucle 30-35, en utilisant le programme Prime (Schrödinger 2011). La géométrie de ce modèle a été optimisée en utilisant les potentiels hybrides MQ/MM. La partie MQ était composée des Cys33, Cys58, Cys102, Asp35 and His101 d'IscU, Cys321 d'IscS et le centre [2Fe-2S], alors que le reste du système était traité en MM.

Afin de voir si le centre [2Fe-2S] pouvait être coordonné par l'Asp35 et la Cys33, l'espace conformationnel de la boucle composée des résidus 30 à 35 a été échantillonné en utilisant l'algorithme implémenté dans Prime, en appliquant des restrictions de distance pour que les Cys33 et Asp35 restent dans la proximité de l'atome de fer du centre, l'autre atome de Fe étant coordonné par les cystéines 58 et 102. Une constante de force de 350 kcal.mol⁻¹.Å⁻² a été utilisée pour restreindre les distances de la liaison Cys33S-Fe et une des liaisons Asp35Oδ-Fe bond distances initialement à 2.3 Å dans un premier temps; les chaînes latérales à 7.5 Å de la boucle 30-35 ont aussi été optimisées. Le modèle le plus bas en énergie issu de l'échantillonnage de boucle avec des valeurs permises dans le diagramme de Ramachandran a alors été optimisé avec les potentiels hybrides MQ/MM. La partie MQ était composée des cystéines 33, 58 et 102, l'His101, le centre [2Fe-2S] et l'Asp35, alors que le reste d'IscU était traité avec la MM.

II.3) Dosage au rose bengal et en fer

II.3.1) Dosage au rose bengal

Il existe plusieurs méthodes pour déterminer la concentration en protéines d'un échantillon : méthode de Lowry, méthode de Bradford, acide bicinchonique (BCA), formule de Warburg et Christian, etc. (Olson and Markwell 2001). La concentration des protéines en solution a été déterminée par dosage colorimétrique au rose bengal (Elliott and Brewer 1978), une méthode peu utilisée. L'intérêt apporté à ce dosage est qu'il est rapide à exécuter et, à la différence de celui de celui de Bradford, fiable pour les protéines à centres FeS. Voici le principe du dosage au rose bengal : la concentration en protéine est déterminée grâce à la

réalisation d'une courbe étalon d'absorbance à 560 nm, en fonction de la concentration connue en protéine standard (Bovine Serum Albumine). Le rose bengal est préparé à 10 mg/mL dans de l'eau.

La BSA est diluée à différentes concentrations connues comprises entre 1 et 80 µg/mL, dans 2 mL de tampon KH_2PO_4 10 mM pH 6.8 (gamme étalon). Il en est de même avec l'échantillon (1 µL à 20 µL d'échantillon à diluer selon la concentration). Un blanc est réalisé avec 2 mL de tampon. Une fois la gamme et les échantillons prêts, 100 µL de rose bengal sont ajoutés, les tubes sont homogénéisés au vortex. Après 2 à 5 min d'attente, 100 µL d'acide acétique glacial à 50 % sont ajoutés et l'absorbance à 560 nm est lue immédiatement après avoir fait le zéro sur le blanc. La courbe de l'absorbance en fonction de la concentration est alors à tracer pour obtenir le coefficient directeur de la droite, qui est le coefficient de proportionnalité entre la concentration et l'absorbance (la droite passant par l'origine). En tenant compte de la dilution, la concentration en protéine de l'échantillon est obtenue.

II.3.2) Dosage en fer

Le dosage en fer se fait selon la méthode de Fish (Fish 1988). Une gamme étalon de 0 à 10 nanomoles de Fe dans un volume final de 65 µL est réalisée avec le sel de Mohr $(\text{NH}_4)_2(\text{SO}_4)_2$ à 500 µM dans de l'eau préparé à partir d'une solution à 10 mM dans du HCl 1N. Les échantillons protéiques sont préparés dans un volume final de 65 µL, de façon à avoir une quantité de fer comprise dans la gamme étalon.

45 µL d'acide perchlorique sont ajoutés et les solutions sont homogénéisées au vortex. Les tubes sont laissés 1 h à température ambiante. Après centrifugation pendant 5 min à 5000 rpm, 90 µL de solution sont prélevés et les solutions suivantes sont ajoutées, dans l'ordre :

- 72 µL de bathophénantroline (1.7 mg/mL dans l'eau)
- 36 µL d'ascorbate de sodium (38 mg/mL dans l'eau) préparée le jour du dosage (car instable)
- 27 µL d'acétate d'ammonium (solution saturée diluée aux 1/3).

Les tubes sont laissés 30 min à température ambiante, et centrifugés pendant 5 min à 5000 rpm. L'absorbance à 535 nm est mesurée dans des cuves en quartz en ayant fait au préalable le zéro sur le blanc. La courbe de l'absorbance en fonction de la concentration est alors à tracer pour obtenir le coefficient directeur de la droite qui est le coefficient de proportionnalité entre le nombre de nanomoles en Fe et l'absorbance (la droite passant par l'origine). En tenant compte du volume de protéine mis en jeu, de sa concentration (la protéine est préalablement dosée au rose bengal) et des dilutions, la concentration en fer de l'échantillon est obtenue.

II.4) Analyse par électrophorèse sur gel de polyacrylamide

La détermination de la composition d'échantillons protéiques a souvent été faite par électrophorèse sur gel de polyacrylamide (PAGE).

Dans le cas du SDS-PAGE, les échantillons ont été traités par le sodium dodécyl sulfate (SDS) et un réducteur (β -mercaptoéthanol) puis chauffés à 98°C, afin de dénaturer la protéine et réduire ses possibles ponts disulfures.

Dans le cas du SDS-PAGE non réducteur, les échantillons ont été traités par le sodium dodécyl sulfate (SDS) puis chauffés à 98°C, afin de dénaturer la protéine.

Dans le cas du gel natif, les conditions ne sont ni dénaturantes ni réductrices.

III) Cristallogénèse et cristallographie des protéines

La structure tridimensionnelle d'une protéine peut être déterminée par cristallographie aux rayons X. La distance interatomique étant de l'ordre de 0.1 nm, ces rayons, qui ont une longueur d'onde comprise entre 0.01 et 10 nm, sont perturbés par les atomes. L'intensité des rayons diffusés par une molécule étant trop faible, cette technique exige d'obtenir des cristaux dans lesquels on retrouve plusieurs fois la protéine empilée de façon périodique dans les 3 dimensions de l'espace. Les ondes diffusées en phase (interférences constructives) vont s'ajouter pour atteindre un signal mesurable sur un détecteur. Le cristal est en quelque sorte un amplificateur du signal.

III.1) Cristallogénèse

Les cristaux de protéine sont obtenus lorsque la protéine passe d'un état liquide à un état solide organisé (Ducruix and Giegé 1999). La technique de cristallisation la plus utilisée est la technique de diffusion de vapeur, soit en goutte assise, soit en goutte suspendue (Figure III-1). Pour la dernière, une goutte de quelques μ L d'un mélange protéine / agent précipitant est déposée sur une lamelle siliconée (pour éviter que la goutte ne s'étale) que l'on retourne sur un puits contenant le même agent précipitant en volume 50 à 1000 fois plus grand que celui de la goutte. La lamelle est scellée au puits avec de la graisse pour rendre le système étanche. La concentration en agent précipitant étant plus petite dans la goutte que dans le puits, des échanges de vapeur (eau ou solvant organique) se font entre la solution réservoir et la goutte afin d'obtenir une concentration en agent précipitant de la goutte approximativement égale à

celle du puits. Dans le cas d'une expérience où le ratio protéine / agent précipitant était de deux au départ, le volume de la goutte à l'équilibre est environ divisé par deux, ce qui implique que la concentration de la protéine double elle aussi. Le volume du puits étant plus grand, sa variation est minime. Si les conditions sont remplies, un phénomène de nucléation peut à ce moment avoir lieu, conduisant à l'apparition de cristaux. Dans le cas de la goutte assise, la goutte est déposée sur un support contenant une cavité siliconée, support qui se trouve dans le puits contenant l'agent précipitant. Le puits est refermé par une lamelle ou du scotch.

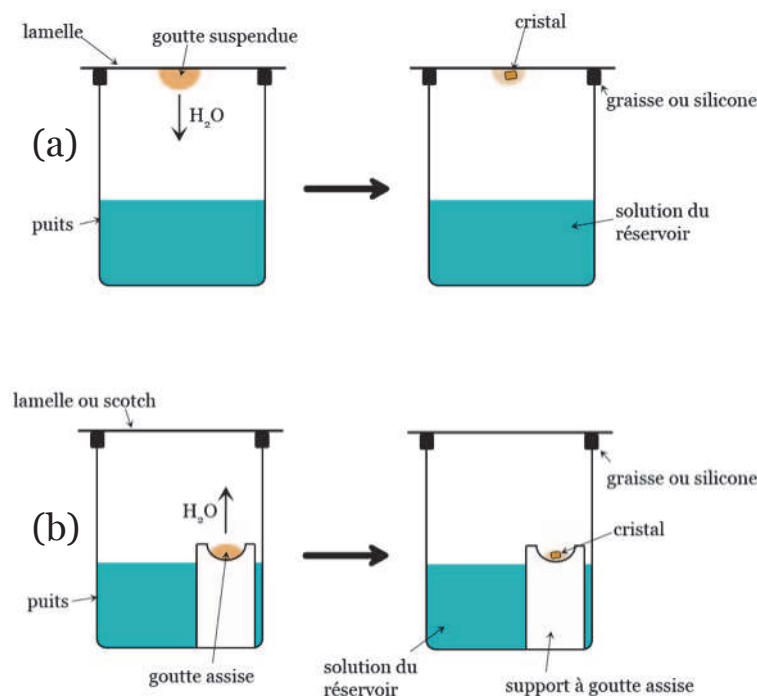


Figure III-1 Technique de la diffusion de vapeur.

(a) en goutte suspendue
(b) en goutte assise

Voici quelques paramètres affectant la cristallisation (et/ou la solubilité de la protéine):

- ✓ sursaturation (concentration de la protéine et l'agent précipitant, voir ci-dessous),
- ✓ température, pH,
- ✓ temps (équibration et de croissance cristalline),
- ✓ force ionique et pureté des produits chimiques (nature du précipitant, tampon, additif),
- ✓ pureté et homogénéité de la protéine.

La solution de cristallisation peut être composée d'un tampon pour fixer un pH, d'un agent précipitant, d'additifs (alcool, sels etc.). Des cofacteurs, inhibiteurs ou détergents peuvent être ajoutés à la protéine.

Les agents précipitants peuvent être de différentes natures :

- ✓ des solvants organiques (isopropanol, MPD, éthanol, etc) qui font diminuer la constante diélectrique du milieu,
- ✓ des sels (ammonium sulfate, NaCl, etc) qui modifient la force ionique et augmentent les interactions hydrophobes entre les molécules de protéine en excluant le solvant,
- ✓ des polymères organiques (PEG, PEGMME, Jeffamine, etc) : la solvation de ces polymères entre en compétition avec la couronne de solvation de la protéine.

On peut décrire un diagramme de phase représentant la solubilité de la protéine en fonction de la concentration en agent précipitant (Figure III-2). La courbe de solubilité sépare 2 zones principales, de sous-saturation et sursaturation. Dans la zone de sous-saturation la concentration en protéine est trop faible pour permettre l'apparition de cristaux. Dans la zone de sursaturation, qui se décompose en zones i) métastable, ii) de nucléation et iii) de précipitation, il y a un excès de protéine qui va soit précipiter, soit cristalliser. Dans la zone de précipitation, comme son nom l'indique, il y a formation de précipité. Dans la zone de nucléation, la sursaturation est suffisante à l'apparition de germes cristallins et à leur croissance en cristaux. La force motrice est le degré de sursaturation qui fournit l'énergie nécessaire à l'association des molécules. Dans la zone métastable, la formation des germes cristallins sera très lente et pourra prendre des mois voire des années (problème de cinétique). Des chemins expérimentaux (Figure III-2) peuvent expliquer l'apparition (chemin i',j',k') ou non (chemin i,j) de cristaux : dans une expérience de diffusion de vapeur, où l'on a mis un volume équivalent de précipitant et de protéine dans la goutte, les concentrations du précipitant et de la protéine sont censées presque doubler ($2C_A$, $2C'_A$, $2C_P$ et $2C'_P$) après équilibration. Pour le chemin i,j, après équilibration le système se trouve dans la zone métastable. Il y a donc très peu de chances pour qu'il y ait apparition de cristaux. Pour le chemin i',j',k', la zone de nucléation est atteinte, il y a apparition de germes, les cristaux vont alors croître. La concentration de protéine en solution décroît jusqu'à atteindre la solubilité, où il y a un équilibre entre la protéine en solution et le cristal.

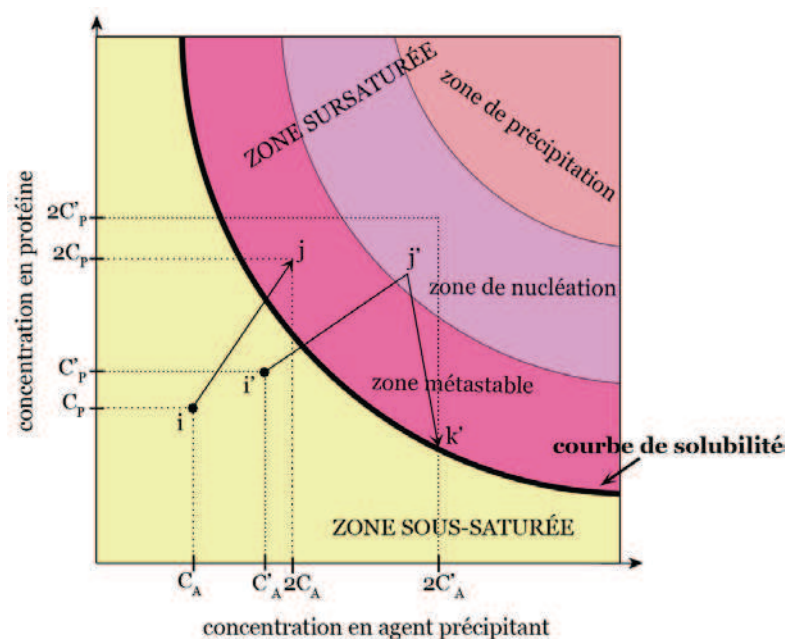


Figure III-2 Diagramme de phase et chemins expérimentaux pour la technique de diffusion de vapeur.

Pour une goutte avec un volume équivalent en protéine et agent précipitant. C_P , C'_P : concentration en protéine ; C_A , C'_A : concentration en agent précipitant. Les points i et i' sont deux conditions initiales en concentration en protéine et agent précipitant.

III.2) Cristallographie aux rayons X

Une fois les cristaux de la protéine d'intérêt obtenus, ils vont être exposés au rayonnement X afin d'obtenir sa structure tridimensionnelle. Lors de l'expérience de diffraction aux rayons X, les cristaux, précédemment congelés, sont gardés sous flux d'azote gazeux à environ 100°K. En effet la température très basse ralentit la propagation du dommage radiatif dans les cristaux, engendré par l'exposition aux rayons X d'haute brillance. Les rayons X éjectent des électrons ce qui forme des radicaux. Ces mêmes électrons vont réduire les centres métalliques et les ponts disulfures. Les radicaux libres, notamment HO^\bullet , diffusent et endommagent les molécules du cristal en induisant d'autres coupures (Burmeister 2000). Ce dommage diminue la qualité de la diffraction. Les cristaux de protéine obtenus sont donc congelés rapidement dans l'azote liquide ou propane liquide en utilisant un cryoprotectant. Les protéines sur lesquelles j'ai travaillé durant ma thèse étant sensibles à l'oxygène, le protocole de congélation spécifique à la boîte à gants a été utilisé (Vernede and Fontecilla-Camps 1999). L'ajout d'un cryoprotectant lors de la congélation évite la formation de glace cristalline à l'extérieur (anneaux de glace sur les clichés de diffraction) et à l'intérieur des cristaux, qui contiennent 30 à 80% d'eau. La formation de cristaux de glace engendre l'augmentation du volume du cristal de protéine ce qui le détériore voire le détruit. Le cryoprotectant (glycérol, MPD, éthylène glycol, PEG 200-600, glucose, etc.) remplace une partie des molécules d'eau du cristal et favorise la transition de l'eau en glace amorphe et non pas en glace cristalline, lorsqu'il est refroidi rapidement dans l'azote liquide ou le propane. Le choix du

cryoprotectant se fait par rapport à la composition de la solution réservoir (si celle-ci contient du PEG on peut utiliser le même PEG en cryoprotection) mais surtout par essais et erreurs.

III.2.1) *La diffraction des rayons X par un cristal*

Lors d'une expérience de diffraction des rayons X par un cristal, ces ondes électromagnétiques, interagissent à travers leur champ électrique avec les électrons des atomes de la protéine. Une partie des ondes est diffusée. La direction du faisceau incident est caractérisée par le vecteur \vec{s}_0 , le faisceau diffusé par le vecteur \vec{s} et on a $\vec{S} = \vec{s} - \vec{s}_0$, vecteur de diffusion ((Janin and Delepierre 1999) et (Drenth 2007)). La diffusion par un élément de volume positionné par le vecteur \vec{r} est décrite en amplitude et en phase par $\delta F = \rho(\vec{r})e^{i2\pi\vec{r}\cdot\vec{S}}\delta V$ avec $\rho(\vec{r})$ densité électronique en tout point M de l'espace et \vec{S} vecteur de diffusion. $\rho(\vec{r})$ est nulle là où il n'y a pas d'atome alors que là où il y a un atome, $\rho(\vec{r})$ est positive et d'autant plus élevée que le numéro atomique (donc le nombre d'électrons) de l'atome est grand. Quand plusieurs éléments δV diffusent à la fois, le facteur de structure $F(\vec{S}) = \iiint_V \rho(\vec{r})e^{i2\pi\vec{r}\cdot\vec{S}}dV$ décrit la structure de l'objet en terme de somme des ondes diffusées par les électrons, avec V étant le volume de celui-ci, $\rho(\vec{r})$ la densité électronique, \vec{r} le vecteur position dans l'espace réel (espace où l'on décrit la matière qui diffuse) et \vec{S} vecteur de diffusion dans un espace réciproque (espace où l'on décrit l'onde diffusée). Le facteur de structure est la transformée de Fourier de la densité électronique. L'intensité est égale au carré $|F(\vec{S})|^2$ de l'amplitude. Pour $\vec{S}=0$, cas où l'onde diffusée est dans la direction de l'onde incidente, $F(0) = \iiint_V \rho(\vec{r})dV$: c'est l'intégrale de la densité électronique dans tout le volume de l'objet diffusant qui est égale au nombre total d'électrons dans l'objet. La phase de $F(0)$ est nulle ce qui signifie que $F(0)$ est l'origine des phases, égale à la phase de l'onde diffusée dans la direction incidente.

Dans un cristal un grand nombre de molécules ont la même orientation, les ondes diffusées peuvent s'ajouter en phase dans certaines directions et atteindre un signal mesurable, il s'agit dans ce cas d'interférences constructives que l'on appelle diffraction. Les points équivalents sont reliés par l'une des translations du cristal. Les ondes peuvent aussi s'annuler ou presque (interférence destructive). Les conditions de diffraction peuvent être décrites par les conditions de Laue ou la loi de Bragg. Selon les conditions de Laue, il y a diffraction si $\vec{r} \cdot \vec{S}$ est un nombre entier ; si ce n'est pas un nombre entier il y a interférence

destructive. $\vec{r} \cdot \vec{S}$ est un nombre entier si $\vec{S} \cdot \vec{a} = h$, $\vec{S} \cdot \vec{b} = k$ et $\vec{S} \cdot \vec{c} = l$; h,k,l entiers et \vec{a} , \vec{b} , \vec{c} paramètres de maille .

Afin de pouvoir représenter les conditions de diffraction, un réseau réciproque est attribué au cristal caractérisé par les vecteurs \vec{a} , \vec{b} et \vec{c} pour son réseau réel. Le réseau réciproque a pour base $\vec{a}^* = \left(\frac{\vec{b} \wedge \vec{c}}{V} \right)$, $\vec{b}^* = \left(\frac{\vec{c} \wedge \vec{a}}{V} \right)$ et $\vec{c}^* = \left(\frac{\vec{a} \wedge \vec{b}}{V} \right)$ avec V le volume de maille cristalline. Le vecteur de diffusion s'écrit alors $\vec{S} = h\vec{a}^* + k\vec{b}^* + l\vec{c}^*$. Il y a diffraction si \vec{S} est un vecteur du réseau réciproque. Les entiers h,k,l sont appelés indices de Miller. On peut représenter les conditions de diffraction par une sphère d'Ewald de rayon $1/\lambda$, dont l'origine est centrée sur le cristal. Lorsque le cristal tourne, le réseau réciproque tourne autour de son origine, et il y a diffraction lorsqu'un des nœuds du réseau réciproque touche la sphère d'Ewald (Figure III–3 (a)). Les taches de diffraction observées correspondent à l'ensemble des directions des vecteurs \vec{s} satisfaisant les conditions de Laue. Selon la direction de diffraction par rapport au réseau du cristal, la sphère d'Ewald intercepte différents nœuds du réseau réciproque. À chaque coïncidence, le vecteur de diffusion \vec{S} est égal à un vecteur \vec{r}^* du réseau réciproque et il se forme une tache de diffraction sur un détecteur dans la direction correspondante.

Dans le cas de la loi de Bragg, il y a diffraction pour l'angle 2θ (angle entre les faisceaux incident et diffracté) vérifiant $n\lambda = 2d_{hkl} \sin \theta$, avec d_{hkl} la distance interréticulaire de la famille de plans sur laquelle se réfléchit le faisceau direct, n un entier et λ la longueur d'onde des rayons X (Figure III–3 (b)).

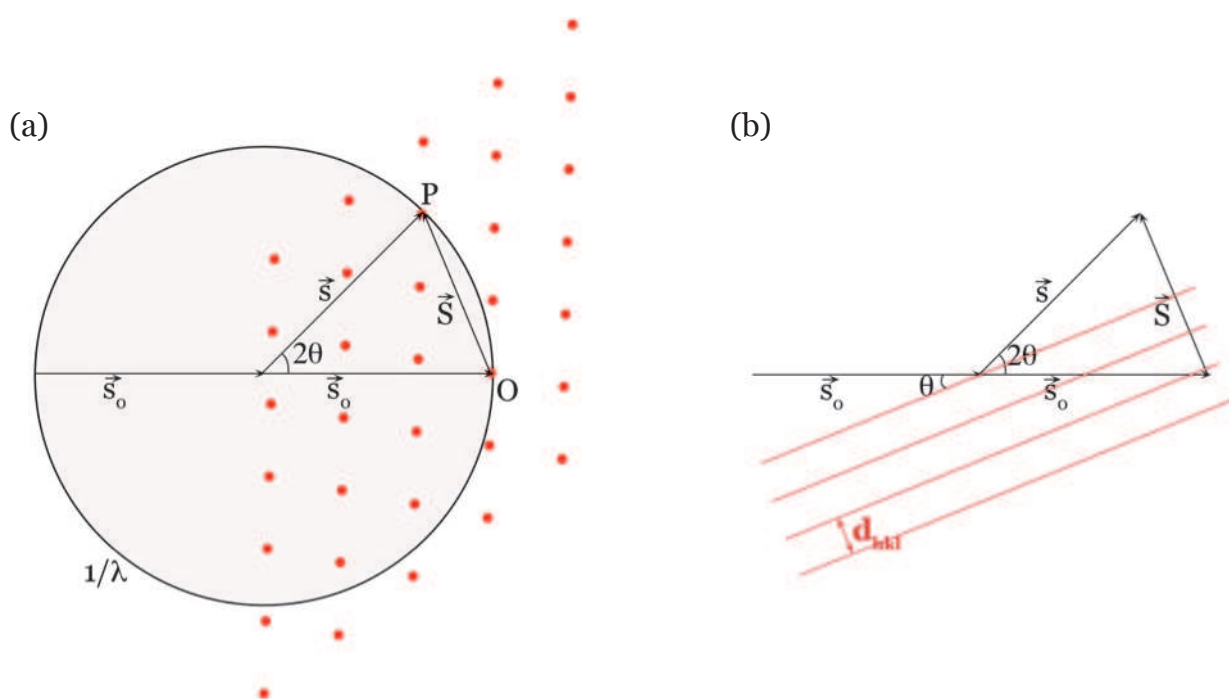


Figure III-3 Phénomène de diffraction selon les lois de Laue (a) et de Bragg (b).

(a) Sphère d'Ewald de rayon $1/\lambda$ et réseau réciproque d'origine O . Le point P touche la sphère d'Ewald, il y a diffraction. (b) Plans réticulaires et la loi de Bragg.

Comme en cristallographie on n'atteint pas en général la résolution suffisante pour visualiser les orbitales atomiques, on considère la densité électronique de chaque atome comme étant sphérique et centro-symétrique. Le facteur de structure peut alors être décrit comme la somme des facteurs de diffusion atomique f_j des différents atomes j auquel est ajouté un terme de déphasage décrivant la position de chaque atome au sein de la structure :

$$\vec{F}_{hkl} = \sum_{j=1}^n f_j e^{i2\pi(hx_j + ky_j + lz_j)} = |F_{hkl}| e^{i\phi_{hkl}} \text{ avec } x_j, y_j, z_j \text{ les coordonnées de l'atome } j \text{ et } \phi_{hkl} \text{ la}$$

phase.

La densité électronique correspond à la transformée de Fourier inverse de la somme des facteurs de structure :

$$\rho(\vec{r}) = \frac{1}{V} \sum_{h,k,l} F_{hkl} e^{-i2\pi(hx + ky + lz)}$$

L'obtention de la densité électronique va nous permettre d'obtenir la structure de la protéine. Lors d'une expérience de diffraction des rayons X les intensités I_{hkl} sont mesurées. L'intensité I_{hkl} est proportionnelle au carré du module du facteur de structure, égal au produit scalaire entre le facteur de structure et son conjugué, le facteur de structure étant un nombre

complexe : $I_{hkl} \propto |F_{hkl}|^2 = \vec{F}_{hkl} \cdot \vec{F}_{hkl}^*$. Nous avons une estimation du module du facteur de structure mais nous n'avons pas d'informations sur la phase qui nous donnerait alors accès à la densité électronique, qui est égale à la transformée de Fourier inverse du facteur de structure. Il s'agit du problème de phase.

III.2.2) *L'enregistrement et le traitement des données de diffraction*

Lors d'une expérience de diffraction des rayons X, le cristal, placé sur un goniomètre, est exposé au faisceau. L'intensité des faisceaux diffractés est enregistrée grâce à un détecteur bidimensionnel sous forme de clichés de diffraction. Le cristal subit une rotation d'en général 1°, angle choisi selon la taille de la maille du cristal, selon un axe parallèle au plan du détecteur et perpendiculaire aux RX, pour couvrir au maximum l'ensemble des directions de diffraction et d'enregistrer la totalité des taches de diffraction non reliées par symétrie.

Les enregistrements de données de diffraction ont été réalisés à l'European Synchrotron Radiation Facility (ESRF) de Grenoble, sur les différentes lignes de lumière dédiées à la biocristallographie (ID23, ID14eh1-4, BM30A etc.). Le synchrotron est un grand instrument dont la fonction est de produire un faisceau de rayons X à partir d'électrons accélérés à une vitesse proche de celle de la lumière. Ce faisceau est de très haute brillance et couvre une large gamme spectrale (énergies de 1.5 à 30 keV à l'ESRF).

Une fois les données enregistrées, elles doivent être traitées. Nous l'avons fait avec le logiciel XDS (Kabsch 2010), couramment utilisé. Ce logiciel est composé de 3 sous-logiciels, XDS, XSCALE et XDSCONV qui vont respectivement indexer/intégrer, mettre à l'échelle et convertir les données dans un format requis par les logiciels de détermination de structure.

L'indexation consiste à rechercher une maille et sa matrice d'orientation permettant d'attribuer à chaque tache de diffraction, enregistrée par le détecteur, ses indices hkl. La détermination de la maille cristalline et l'analyse des absences systématiques permettent de sélectionner le groupe d'espace le plus probable. La position de chaque tache étant connue grâce à la maille et la matrice d'orientation, l'intensité de ces taches est alors mesurée pour chaque image. XDS a la particularité de calculer un profil tridimensionnel des taches de diffraction.

Une même réflexion ou des réflexions équivalentes par symétrie devraient avoir la même intensité sur les différentes images. Cependant, plusieurs paramètres font que les réflexions

équivalentes n'auront pas la même intensité. Par exemple, au fur et à mesure de l'enregistrement de données, le cristal subit des dégradations dues aux rayons X, qui diminuent l'intensité des taches diffractées. Lors d'enregistrement de jeux de données à l'ESRF, les réflexions équivalentes ne vont pas recevoir exactement la même intensité de rayons X si l'expérience prend plusieurs heures, à cause de la décroissance de l'intensité du faisceau incident, leurs taches de diffraction n'auront donc forcément pas les mêmes intensités. La forme du cristal peut également entraîner des variations de l'intensité des taches de diffraction au cours de la collecte en fonction de l'orientation du cristal, selon qu'un volume plus ou moins important est traversé par le faisceau de rayons X. Il est donc important de faire des corrections permettant de tenir compte de ces phénomènes. Il s'agit de la mise à l'échelle des données, qui vont minimiser globalement l'écart entre les intensités des réflexions équivalentes. Après mise à l'échelle, les intensités moyennes peuvent être calculées à partir des mesures d'intensités des taches équivalentes.

Une fois cette correction effectuée, le nombre de réflexions est réduit en moyennant les intensités des réflexions équivalentes. Dans une expérience classique, les réflexions d'indices h, k, l sont considérées, par la loi de Friedel, comme étant équivalentes aux réflexions d'indices $-h, -k, -l$, faisant encore augmenter la redondance.

Les amplitudes F des facteurs de structures sont alors calculées à partir des intensités I . Le fichier de sortie (fichier mtz) contient la liste des indices h,k,l , FP (valeur de F native) et $SIGFP$ (sigma de FP).

III.2.3) Qualité d'un jeu de données

Différents critères permettent d'évaluer la qualité d'un jeu de données :

- La résolution limite (en Å), qui est liée au niveau des détails visibles dans les cartes de densité électronique. La limite de résolution peut être déterminée par l'expérience ou par la qualité de diffraction du cristal. Nous avons fixé cette limite pour un $R_{\text{sym}} \leq 40\%$ ou un rapport $I/\sigma(I) \geq 2$ dans la dernière tranche de résolution ;
- la redondance donne le nombre moyen de mesures de chaque réflexion équivalente. Plus des équivalents d'une réflexion seront mesurés, moins il y aura d'erreur de mesure;

- la complétude représente le pourcentage de réflexions uniques mesurées par rapport au nombre théorique de réflexions uniques à mesurer pour une résolution donnée. Elle doit être supérieure à 90% ;
- le rapport signal sur bruit $I/\sigma(I)$ est le rapport entre l'intensité moyenne des réflexions mesurées et l'écart type de ces intensités. Ce rapport diminue avec l'augmentation de la résolution ;

- le facteur R_{sym} (en %) : $R_{\text{sym}} = \frac{\sum_{hkl} |I_{hkl} - \langle I \rangle|}{\sum_{hkl} I_{hkl}}$, permet de comparer l'intensité de

chaque réflexion équivalente I_{hkl} à la valeur moyenne $\langle I \rangle$. Plus sa valeur est basse, meilleure est la cohérence du jeu de données.

III.2.4) Estimation du nombre de molécules par unité asymétrique

L'unité asymétrique est le plus petit volume de la maille contenant des points non reliés entre eux par la symétrie cristallographique. Elle peut contenir plusieurs molécules qui sont alors reliées par des opérateurs de symétrie, que l'on appelle symétrie non cristallographique (NCS). Le nombre de molécules par unité asymétrique s'obtient par le calcul des coefficients de Matthews. Ce dernier (Matthews 1968) a montré que pour la plupart des cristaux de

protéine, le rapport V_m : $V_m = \frac{V}{M \times n \times Z_a}$ entre le volume de la maille (V) et le poids

moléculaire contenu dans celle-ci, était compris entre environ 2 et 4 Å³/Da (Kantardjieff and Rupp 2003). Le poids moléculaire contenu dans la maille est le produit de la masse moléculaire de la protéine M , du nombre n d'unités asymétriques dans la maille et du nombre de molécules dans l'unité asymétrique Z_a . La portion de solvant (V_s) contenu dans le cristal

se calcule par la formule $V_s = 1 - \frac{1.23}{V_m}$.

III.2.5) Remplacement moléculaire

Après collecte des données cristallographiques, il faut alors résoudre le problème de la phase mentionné précédemment. Pour cela il existe trois méthodes classiques (Taylor 2010):

1) le remplacement isomorphe, utiliser la propriété de diffusion anormale et 3) le remplacement moléculaire. Les structures résolues durant ma thèse l'ont été par ce dernier.

La technique du remplacement moléculaire consiste à utiliser un modèle de départ qui est une structure protéique connue dont la séquence est fortement similaire avec la protéine étudiée. La résolution des structures des protéines HmdB, IscS et IscS-IscU a été faite par remplacement moléculaire, disposant de modèles possédant de fortes similitudes de séquence pour ces structures. Les six paramètres de la transformation géométrique permettant de passer du modèle à la protéine vont être recherchés (trois pour la rotation et trois pour la translation). La modèle est d'abord placé dans la maille par une fonction de **rotation** et une fois que les orientations favorables ont été trouvées, elles sont soumises à une fonction de **translation**, qui essaye de placer la (ou les) molécule(s) correctement orientée(s) dans la maille. La décomposition en deux étapes permet une recherche de la solution de remplacement moléculaire dans deux espaces indépendants, de trois paramètres chacun, à la place d'un espace à six paramètres. Le nombre de calculs de corrélation nécessaire à la recherche est ainsi diminué, diminuant le temps de calcul. Le meilleur placement du modèle dans le cristal est obtenu lorsque le facteur de structure du modèle F_{calc} est proche du facteur de structure des données observées F_{obs} (dont les phases ont été estimées avec le modèle). Le logiciel que nous avons utilisé est PHASER (McCoy, Grosse-Kunstleve et al. 2007). Dans PHASER, la recherche des solutions, c'est-à-dire des meilleures corrélations entre F_{obs} et F_{calc} , se fait en utilisant la méthode du maximum de vraisemblance (Maximum Likelihood) (Evans and McCoy 2008). La méthode consiste à construire une expression analytique de vraisemblance de F_{obs} en fonction de F_{calc} , permettant de trouver la valeur la plus vraisemblable de la phase. La recherche s'effectue dans l'espace réciproque. Pour chaque orientation et position du modèle F_{calc} dans le cristal, la probabilité d'obtenir les amplitudes expérimentales $|F_{\text{obs}}|$ est calculée.

En mode « automatique », le programme PHASER fait le remplacement moléculaire en 6 étapes :

- Correction d'anisotropie : les cristaux de macromolécules biologiques ne diffractent pas de manière homogène car, d'une part les atomes équivalents n'ont pas exactement la même position d'une maille à l'autre, d'autre part l'agitation thermique fait vibrer les atomes autour de leur position d'équilibre. Ce désordre est quantifié par le facteur de température $B = 8\pi^2 \langle (\Delta r)^2 \rangle$, avec $\langle (\Delta r)^2 \rangle$ le déplacement atomique moyen. L'intensité des réflexions peut être exprimée en fonction de B : $I_{hkl} = I_0 \exp(-B \sin^2 \theta / \lambda^2)$, avec I_{hkl} l'intensité de la réflexion, θ l'angle de diffraction et λ longueur d'onde du rayon X incident. L'intensité des ondes diffractées diminue donc avec le facteur de température B et avec la résolution. Il est nécessaire de supprimer les variations

anisotropiques avant le remplacement moléculaire en normalisant les facteurs B de chaque réflexion.

- Génération d'un modèle (« ensembling ») : dans un premier temps, le nombre de molécules à placer dans la maille est estimé selon le coefficient de Matthews V_M . Les facteurs de structure du modèle sont ensuite calculés (F_{calc}) pour être comparés aux données expérimentales. Dans PHASER, ce calcul prend en compte l'écart-type moyen (ou rmsd) entre la structure du modèle et la structure inconnue. Le rmsd en C α peut être estimé en fonction du pourcentage d'identité de séquence par la formule suivante : $rmsd = 0,4 \exp(1,87 \times (1 - \% \text{ d'identité de séquence}))$.
- Fonction de rotation R : la recherche par rotation selon les angles Euler est guidée par la plus haute vraisemblance sur les intensités, donc les amplitudes. Une première recherche rapide est faite par la fonction LERF (Likelihood-Enhanced fast Rotation Function) qui est une simplification de la fonction de rotation. Les plus hautes valeurs du LERF sont ensuite sélectionnées pour être re-calculées dans la fonction de rotation complète MLRF (Maximum-Likelihood Rotation Function) qui permet une meilleure discrimination.
- Fonction de translation T : suite à l'orientation, le positionnement du modèle est réalisé selon les symétries du cristal, avec la même stratégie de calcul : LETF (Likelihood-Enhanced fast Translation Function), puis MLTF (Maximum-Likelihood Translation Function). La recherche est guidée par la plus haute vraisemblance des phases cette fois.
- Fonction d'empilement : l'empilement cristallin des solutions proposées est évalué sur la proximité géométrique des C α inter-moléculaires, avec une valeur limite fixée à 2 Å. Le nombre d'excès autorisé doit être généralement nul, à moins que l'identité de séquence avec le modèle soit faible ou qu'il y ait potentiellement de larges boucles flexibles.
- Affinement en corps rigide : la solution est une molécule modèle calculée (F_{calc}) dont les coordonnées doivent être finement remaniées pour refléter au mieux les données expérimentales (F_{obs}) : c'est le rôle de l'affinement cristallographique. À l'issue du remplacement moléculaire, PHASER effectue un premier affinement en corps rigide afin d'optimiser la position du modèle dans le cristal.

À chaque solution de rotation ou translation proposée, des indices de qualité sont donnés : le LLG et le Z-score. Le Z-score (RFZ pour la fonction de rotation et TFZ pour la fonction de translation) correspond à la différence entre la solution sélectionnée et la moyenne. Lorsque

ce score est inférieur à 5, il s'agit d'une solution incorrecte ; entre 5 et 6, elle est improbable ; entre 6 et 7, elle est possible ; entre 7 et 8, elle est probable et au-delà de 8, elle est définitive. Le LLG (Log Likelihood Gain) est une indication de la qualité de la solution quand on la compare à une solution au hasard. Plus le LLG est élevé, meilleure est la solution.

III.2.6) *La diffusion anormale*

Le phénomène de diffusion anormale est dû au fait que le facteur de diffusion atomique (f) d'un atome varie fortement, en module et en phase, dans la région du seuil d'absorption de l'atome. Dans le cas d'électrons libres, les facteurs de structure respectent la loi de Friedel : $|F_{hkl}| = |F_{-h-k-l}|$ et $\Phi_{hkl} = \Phi_{-h-k-l}$. Lorsque des atomes diffusent normalement, le rayonnement diffusé est déphasé de 180° par rapport au rayonnement incident. Dans ce cas, le facteur de diffusion f est réel. Lorsque la longueur d'onde utilisée correspond au seuil d'absorption d'un atome lourd dans le cristal, il y a alors absorption du rayonnement incident par les électrons des couches profondes de l'atome lourd. Ces électrons sont alors excités et entrent en résonance avec les photons. Le rayonnement diffracté par le diffuseur anormal est déphasé de 90° au lieu de 180° . Pour tenir compte de ce possible déphasage le facteur de diffusion f est représenté par un nombre complexe qui dépend de la longueur d'onde λ du rayonnement incident : $f = f^0 + f'(\lambda) + if''(\lambda)$. f^0 est le facteur de diffusion réel en absence de diffusion anormale ; f' (dépendant de λ) est la partie réelle du facteur de diffusion anormale due à la contribution des électrons des couches profondes ; f'' (dépendant de λ) est la partie imaginaire du facteur de diffusion anormale due à l'absorption des RX par l'électron qui entre en résonance et réémet le photon avec un déphasage de 90° . Lorsque le rayonnement incident est loin du seuil d'absorption de l'atome f' et f'' sont négligeables. Si la longueur d'onde utilisée est proche du seuil d'absorption d'un atome, f' et f'' prennent des valeurs significatives. Le déphasage, dû au diffuseur anormal, brise la loi de Friedel. $|F_{hkl}| \neq |F_{-h-k-l}|$ et $\Phi_{hkl} \neq \Phi_{-h-k-l}$. La représentation des vecteurs correspondants dans un diagramme d'Argand montre bien cet aspect (Figure III-4).

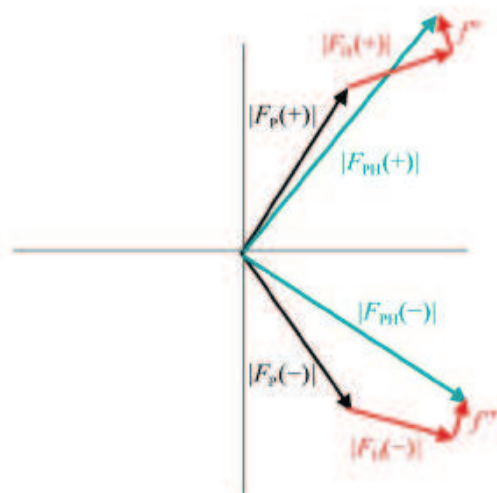


Figure III-4 Loi de Friedel fausse en présence de diffusion anormale.

$f = f^0 + f'(\lambda) + if''(\lambda)$. $|F_{hkl}| \neq |F_{-h-k-l}|$ ou $|F_{PH}(+)| \neq |F_{PH}(-)|$.

(+)=indices h,k,l ; (-)=indices $-h,-k,-l$;

F_p : facteur de structure des diffuseurs normaux ;

F_H : partie réelle ($f_0 + f'$) du facteur de diffusion du diffuseur anormal ;

f'' : partie imaginaire du facteur de diffusion du diffuseur anormal.

Figure tirée de Taylor (Taylor 2010).

III.2.7) La résolution de la structure de la protéine

III.2.7.1) La modification de densité

Les premières cartes de densité électronique calculées peuvent ne pas être interprétables. Il est alors possible d'introduire quelques contraintes générales sur la densité électronique, de manière à améliorer les phases expérimentales. Ce sont les étapes de modification de densité électronique, comme par exemple la méthode d'aplatissement de solvant ou de moyennation dans le cas où il existe des symétries non cristallographiques. Toutes ces méthodes permettent d'améliorer les phases expérimentales car elles introduisent des relations entre les différents facteurs de structure, et donc des contraintes sur les phases.

L'aplatissement de solvant :

Cette procédure est basée sur le fait que le solvant, non ordonné dans le cristal, ne participe pas à la diffraction, sauf à basse résolution. La densité électronique correspondant aux zones de solvant doit donc être plate, ce qui n'est pas le cas, du fait des erreurs de phase. La méthode décrite par Wang (Wang 1985) consiste à discriminer les zones de solvant des zones de protéine, sur la base de la valeur locale moyenne de la densité électronique, en tenant compte du pourcentage de solvant du cristal pour définir le seuil de coupure. Cela permet de déterminer un masque ayant pour valeur 1 pour les zones de protéine et 0 pour les zones de solvant. Il suffit alors d'appliquer ce masque, puis de recalculer des phases à partir de cette nouvelle carte de densité électronique nivelée. Ensuite, il est possible de refaire quelques cycles de détermination de masque et nivellement jusqu'à convergence du facteur R_{free} . Ce facteur R_{free} est le même que celui utilisé pour suivre l'affinement (voir paragraphe suivant). Il

mesure l'accord entre des F_{obs} mis de côté, et n'intervenant pas dans les calculs de carte de densité électronique, et les F_{calc} de la procédure. Cette méthode sera d'autant plus puissante que le pourcentage de solvant dans le cristal est élevé.

La moyennation - symétries non cristallographiques :

L'existence de symétries non cristallographiques permet de remplacer la densité électronique de chaque molécule de l'unité asymétrique par la densité électronique moyenne correspondante. Cette procédure est très efficace dans l'amélioration des phases. Elle nécessite la détermination préalable des opérateurs de symétrie non cristallographique reliant les différents motifs de l'unité asymétrique, ainsi que le masque décrivant la zone correspondant à la molécule concernée par la symétrie non cristallographique.

Procédure itérative de modification de densité électronique (Figure III-5):

L'étape de modification de densité est une procédure itérative. A chaque cycle, de nouvelles phases calculées sont obtenues et combinées avec les phases expérimentales. Cela est effectué jusqu'à convergence du facteur R_{free} . Normalement, il est possible d'obtenir une carte de densité électronique de meilleure qualité que la carte expérimentale du départ, surtout s'il a été possible de réaliser une moyennation sur plusieurs molécules de l'unité asymétrique.

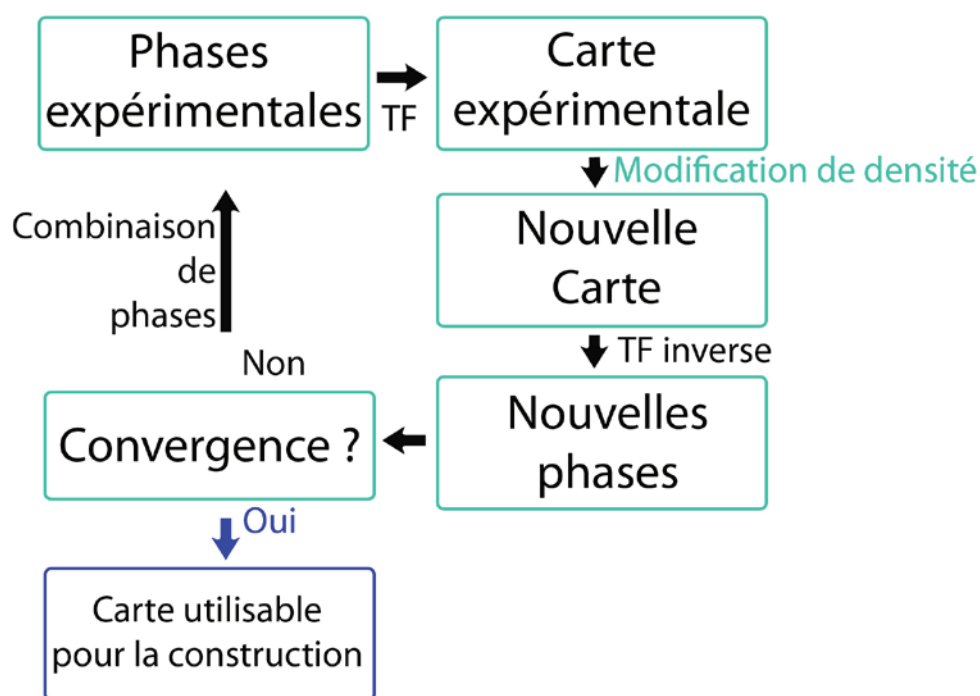


Figure III-5 Procédure de modification de densité électronique.

TF : Transformée de Fourier.

III.2.7.2)

La construction et l'affinement de la structure

Lorsqu'une première estimation des phases a été obtenue par une des méthodes de phasage, une carte de densité électronique initiale de la structure de la protéine peut être obtenue. Les facteurs de structure calculés à partir de ce modèle initial ne sont généralement que partiellement en accord avec les facteurs de structure observés. Une mesure du désaccord existant entre les facteurs de structure calculés (F_{calc}) à partir du modèle et les facteurs de structure observés (F_{obs}) issus des données expérimentales est représentée par le facteur R (ou

$$R_{work}) : R = \frac{\sum_{hkl} \|F_{obs} - F_{calc}\|}{\sum_{hkl} |F_{obs}|}.$$

A l'issue de l'étape de phasage, le facteur R est en général de l'ordre de 40%. Le processus d'affinement consiste alors à améliorer le modèle de manière à trouver un meilleur accord entre les facteurs de structure calculés et les facteurs de structure observés pour atteindre des valeurs de R allant de 10 à 20 %, en fonction de la qualité des données de diffraction. L'ajustement du modèle consiste à affiner à la fois la position (les coordonnées atomiques x, y et z), le facteur de température (facteur B) et parfois le facteur d'occupation de chacun des atomes de la molécule. En raison de leur faible interaction avec les rayons X, les atomes d'hydrogène, qui ne possèdent de fait qu'un seul électron, ne sont pas modélisés dans les structures cristallographiques, à part à très haute résolution. Le nombre d'atomes d'une protéine est important et donc le nombre de paramètres à affiner (4 au minimum par atomes) l'est encore plus. C'est la raison pour laquelle est inclus dans le processus d'affinement un ensemble de contraintes géométriques, telles que la longueur des liaisons interatomiques, les angles de valence, issues de la connaissance de la structure de petites molécules ou de protéines à haute résolution. En pratique, le processus d'affinement cristallographique fait intervenir, tour à tour, des phases de reconstruction manuelle du modèle à l'aide de logiciels de visualisation (tel que le logiciel graphique COOT (Emsley, Lohkamp et al. 2010)) et des cycles d'affinements numériques des paramètres de celui-ci (coordonnées et facteur d'agitation thermique). La reconstruction manuelle permet de modifier le modèle, de manière importante parfois, en le positionnant correctement dans les cartes de densité électronique et l'affinement numérique permet ensuite de compenser les erreurs introduites lors de la reconstruction manuelle. Au cours de cette procédure d'affinement, la "pseudo-énergie potentielle" du système (E_{tot}) est minimisée. Le programme d'affinement REFMAC5 (Murshudov, Vagin et al. 1997) utilise la méthode statistique du maximum de vraisemblance

en tant que critère de convergence, dont le concept central repose sur la définition de la vraisemblance du modèle (des facteurs de structure calculés) comme étant la probabilité de réaliser la mesure des facteurs de structure observés, si l'hypothèse du modèle est juste. L'affinement se fait alors par la recherche du maximum du logarithme de la probabilité. Le logiciel AUTOBUSTER contenant BUSTER-TNT (Blanc, Roversi et al. 2004), procède lui aussi par maximum de vraisemblance.

III.2.7.3) Les critères de suivi de l'affinement

Afin d'évaluer et de valider la bonne progression de l'affinement ainsi que la qualité du modèle final, plusieurs critères sont à la disposition de l'expérimentateur. Le facteur R, déjà présenté, permet de juger de l'accord entre le modèle (facteurs de structure calculés) et les facteurs de structure mesurés. Néanmoins, les méthodes d'affinement numérique ayant pour but de minimiser au mieux ce facteur, il peut arriver que les erreurs du modèle compensent les erreurs des mesures. De plus, au cours des étapes de reconstruction manuelle du modèle, l'interprétation des cartes de densité présente une certaine part de subjectivité qui peut conduire l'expérimentateur à introduire des erreurs. C'est pourquoi l'utilisation d'un autre facteur permettant une estimation moins biaisée de la qualité du modèle a été introduite, en tant que co-critère d'évaluation. Ce facteur, appelé facteur R_{free} (Brunger 1992), est calculé de la même manière que le facteur R mais à partir d'un ensemble de réflexions choisies aléatoirement (5 à 10% des données) et non incluses dans l'affinement, appelé jeu de données test. La valeur de ce R_{free} mesure ainsi la justesse avec laquelle le modèle permet de prédire les données de diffraction exclues de l'affinement. En effet, les propriétés des données de diffraction sont telles que chaque réflexion contient des informations concernant l'ensemble des atomes de la structure. C'est pourquoi des modifications apportées au modèle qui ne permettent pas d'améliorer l'accord entre ce dernier et les données de diffraction, n'améliore pas non plus la capacité du modèle à être en accord avec les réflexions omises de l'affinement. Dans ce cas, la valeur du R_{free} reste constante, ou augmente même. L'écart moyen entre ces deux facteurs est de 7 %. De nombreux facteurs tels que la qualité des données, la procédure d'affinement (utilisation de contraintes géométriques plus ou moins strictes ou de contraintes de symétrie non cristallographique), la complétude du modèle ou la présence d'erreurs influent directement sur la différence entre ces deux facteurs qui dans tous les cas doit être la plus faible possible.

D'autres critères, basés cette fois sur la géométrie, permettent également de valider un modèle. Ainsi, les déviations moyennes des longueurs et angles de liaisons entre atomes du

modèle par rapport aux valeurs standards permettent d'estimer la qualité stéréochimique de celui-ci (root mean square deviation, r.m.s.d). La qualité du modèle est jugée correcte lorsque ces déviations ne sont pas supérieures aux variations physiques déjà observées dans les structures de petites molécules (environ 0.02 Å pour les longueurs et 2Å pour les angles). Le diagramme de Ramachandran (Ramachandran, Ramakrishnan et al. 1963), lui aussi, permet de juger de la qualité stéréochimique du modèle. En effet, en raison des contraintes d'encombrement stérique dues aux chaînes latérales des résidus d'acides aminés, les liaisons peptidiques ne peuvent adopter qu'un nombre restreint de conformations énergétiquement favorables. Ces conformations, caractérisées par les angles de torsion ϕ et ψ le long de la chaîne principale, se concentrent dans des zones préférentielles du diagramme de Ramachandran (ϕ en fonction de ψ). Ces zones sont directement reliées aux structures secondaires (hélices α , feuillets β) rencontrées dans les structures de protéines. Les chaînes latérales, elles aussi, adoptent un nombre limité de conformations préférentielles regroupées sous le terme de rotamères qui permettent de valider leur conformation dans le modèle. Enfin, les contacts atomiques comme les interactions hydrophobes ou la géométrie des liaisons hydrogène potentielles sont encore autant d'informations pour juger de la qualité du modèle cristallographique.

III.2.7.4)

Les cartes de densité électronique

La densité électronique obtenue dessine l'enveloppe de la molécule. On peut en effet considérer que le nuage électronique entourant les atomes matérialise leur surface de Van der Waals, le noyau des atomes étant modélisé au pic de densité électronique. Lors de la reconstruction graphique de la molécule, tout comme pour l'interprétation finale, le modèle est visualisé dans des cartes de densité électronique. Le contour de la molécule est matérialisé par la carte de type $2F_{\text{obs}} - F_{\text{calc}}$, le modèle est construit à l'intérieur de cette enveloppe en respectant des contraintes géométriques. Sont également utilisées des cartes de densité de type $F_{\text{obs}} - F_{\text{calc}}$, ou cartes de Fourier différence. Celles-ci mettent en évidence les erreurs du modèle : les pics positifs indiquent les zones où il manque des électrons, et les pics négatifs les zones où les électrons sont en excès. Il est important de noter que les cartes de densité électronique, notamment de type $2F_{\text{obs}} - F_{\text{calc}}$, peuvent être en partie biaisées par l'inexactitude du modèle. Afin d'étudier une région particulière de la structure (ligand par exemple) en s'affranchissant totalement du biais introduit par le modèle, il est également possible de calculer des cartes de densité électronique de type "omit". En pratique, les atomes de cette région du modèle sont éliminés, puis après quelques cycles d'affinement pour retirer un éventuel biais, des cartes

sont recalculées, le reste du modèle (soit la majeure partie) apportant l'information de phase pour cette région d'intérêt. Les cartes ainsi obtenues ne présentent plus de biais dû au modèle dans cette région particulière.

Les cartes de différence anormale permettent de visualiser et localiser la diffusion anormale d'un atome. Pour cela une carte avec la différence ($F(+)$ - $F(-)$, DANO dans XDS) (où $F(+)$ et $F(-)$ sont les paires de Friedel non équivalentes dans le cas de la diffusion anormale) est calculée.

IV) Spectroscopies

IV.1) Séquençage N-terminal de protéine

La détermination de séquence en N-terminal de protéine a été faite par Jean-Pierre Andrieu, à l'IBS. Les bandes protéiques sur gel de polyacrylamide ont été transférées sur une membrane de polymère du fluorure de vinylidène (PVDF), pour être isolées une à une en les découpant. La détermination de la séquence N-terminale est basée sur la dégradation d'Edman et réalisée sur un séquenceur Applied Biosystems modèle 492 (N° de série 9510287J). Les acides aminés sous forme phenylthiohydantoine générés à chaque cycle sont identifiés par HPLC sur Applied Biosystems Model 140C, et grâce à une analyse des données par le logiciel Applied Biosystems Model 610A (version 2.1). Les procédures et réactifs sont ceux préconisés par le constructeur.

IV.2) Spectroscopie UV-visible

L'absorption d'un photon peut permettre la transition d'un état électronique fondamental S_0 à un état électronique excité S_1 par passage d'un électron d'une orbitale à une autre. L'énergie d'un photon absorbé lors d'une transition entre deux niveaux d'énergie est donnée par la relation $E = (hc)/\lambda$, avec λ la longueur d'onde, h la constante de Planck et c la célérité). Différentes énergies vont permettre différentes transitions et dans le cas des protéines, certaines de ces transitions sont permises avec l'absorption de photons proche de la lumière UV et visible. Les résidus aromatiques des protéines absorbent à une longueur d'onde de 280 nm et dans les protéines à centres FeS, le transfert de charge du soufre vers le fer provoque une absorption comprise entre 300 et 650 nm. Cette valeur varie en fonction du type d'agrégat fer-soufre.

Les spectres UV-visible ont été enregistrés dans des cuves en quartz avec le spectrophotomètre Cary 50 Bio de Varian (qui fait maintenant partie de Agilent Technologies). Dans le cas de spectres en condition anaérobie, des caoutchoucs ont été placés sur les cuves puis recouvertes de parafilm, afin de rendre les cuves étanches.

IV.3) Spectroscopie de résonance paramagnétique électronique (RPE)

L'analyse d'échantillons du complexe IscSU par spectroscopie RPE a été effectuée par le Dr Geneviève Blondin au LCBM du CEA Grenoble.

La spectroscopie de résonance paramagnétique électronique (RPE) permet d'étudier les espèces dites paramagnétiques qui contiennent un ou plusieurs électrons non appariés. L'électron possède un spin électronique : $S = 1/2$. Dans le cas où un atome est entouré de plusieurs électrons, ils vont s'apparier deux à deux (états de spin $m_s = +1/2$ ou $-1/2$). Le spin total (S) de la molécule sera alors nul. Mais lorsque ce n'est pas le cas, le spin de la molécule sera un multiple de $1/2$, et la RPE peut s'appliquer ($S \neq 0$). Cette dernière situation se rencontre avec les radicaux libres (un électron non apparié) ou avec des molécules incorporant des atomes métalliques (un ou plusieurs électrons non appariés selon la nature du ou des atomes métalliques impliqués). L'entité magnétique est placée dans un champ magnétique (H) variable et reçoit une onde électromagnétique d'énergie $(hc)/\lambda$ dont la fréquence se situe dans la région des micro-ondes. Les deux états de spin m_s déterminent alors deux niveaux d'énergie. Le phénomène de RPE se produit lorsque l'énergie de l'onde $(hc)/\lambda$ est égale à $g\beta H$, l'énergie de transition entre deux états de spin m_s de l'électron (g : facteur de Landé ($g_e=2.0023$ pour l'électron libre) et β : magnéton de Bohr). Dans ces conditions l'onde est absorbée par le système et la transition entre les deux états de spin m_s a lieu. Le facteur g diffère de la valeur g_e en fonction de la nature du centre paramagnétique et de son environnement moléculaire. En général, g est anisotrope et se définit par les trois composantes g_x , g_y et g_z . Ces trois composantes correspondent aux points d'inflexions et à l'extremum du spectre RPE enregistré au cours du balayage des valeurs du champ magnétique. Le spectre RPE obtenu expérimentalement est la dérivée du signal d'absorption, afin d'obtenir plus facilement les trois valeurs.

Pour l'étude des protéines à centres FeS, la RPE nous informe sur le type de centre, son état redox et sa quantité puisque le signal peut être intégré et la valeur de l'intégration comparée à celle d'une solution standard de concentration connue.

Les échantillons à analyser (85 à 100 μL , concentration en complexe IscSU de 300 μM) ont été transférés dans des tubes pour le spectromètre RPE, en boîte à gants en condition anaérobie, et ont été congelés dans un bloc qui avait été préalablement refroidi dans l'azote liquide à l'extérieur de la BAG.

IV.4) Spectroscopie Raman de résonance

Les expériences concernant le complexe IscSU ont été réalisées au LEPMI à Grenoble par le Dr Thierry Pagnier.

Un faisceau de lumière monochromatique produit par un laser continu de fréquence ν_0 , est focalisé sur l'échantillon à analyser. L'énergie E_0 de la radiation incidente est définie par la relation $E_0 = h\nu_0 = (hc)/\lambda_0$, avec λ_0 la longueur d'onde choisie, ν_0 la fréquence, h la constante de Planck et c la célérité. Les photons incidents sont en grande partie transmis, réfléchis ou absorbés et une fraction bien plus faible est diffusée : une partie de la lumière est diffusée à la même fréquence ν_0 que la radiation incidente, c'est la diffusion Rayleigh (diffusion élastique), une partie infime des photons incidents est diffusée avec changement de fréquence : c'est la diffusion Raman (diffusion inélastique) : elle représente environ 1 photon Raman pour 10^8 photons incidents. La diffusion Raman correspond à l'interaction de la polarisabilité de la molécule avec la radiation incidente, entraînant un moment dipolaire induit dans la molécule et l'émission d'un photon Raman. Si la fréquence $\nu_0 - \nu$ du photon diffusé est inférieure à celle du photon incident (ν_0), il y a gain d'énergie vibrationnelle pour la molécule (et perte d'énergie pour le photon) : on parle de raie Stokes. Si au contraire le photon incident est diffusé à une fréquence $\nu_0 + \nu$, il y a perte d'énergie vibrationnelle pour la molécule (et gain d'énergie pour le photon) : on parle de raie anti-Stokes. L'intensité de la diffusion anti-Stokes décroît rapidement lorsque ν augmente car la population du niveau excité diminue conformément à la loi de Boltzmann. C'est pourquoi en diffusion Raman on ne considérera que les raies Stokes, plus intenses. Un spectre Raman présente l'intensité Raman diffusée en fonction de la différence de fréquence entre photon incident et photon diffusé : si l'on considère les raies Stokes, cette différence est positive, et nulle dans le cas de la diffusion Rayleigh. En pratique, il est d'usage de convertir cette différence de fréquence entre photon incident et photon diffusé en nombre d'onde. Le nombre d'onde k (en cm^{-1}) est défini par la relation suivante : $k = 1/\lambda = \nu/c$.

Pour les protéines à centres FeS, il est possible mesurer les fréquences de modes de vibrations associés aux élongations des liaisons Fe-S (situées entre 200 et 450 cm^{-1}) en

employant une onde incidente dont la longueur d'onde est dans la bande d'absorption du chromophore. Les fréquences des modes de vibration observés donnent des informations sur la géométrie globale du centre et sur ses ligands.

Les expériences ont été réalisées à température ambiante et à une longueur d'onde de 514 nm. Les échantillons contenant des cristaux du complexe ont été préparés en boîte à gants en condition anaérobie, les lamelles contenant les gouttes de cristallisation ont été recouvertes d'une lamelle ne contenant pas de métaux. De la graisse avait déposée entre les deux lamelles, afin d'avoir une étanchéité.

Résultats et discussion

I) Etude structurale de la prot ine HmdB

La prot ine Hmd (ou HmdA : H₂ forming m thyle-H₄-methanopterin d hydrogenase), pr sente uniquement chez certains microorganismes arch es m thanog nes, catalyse la r action suivante (Figure I-1) (Schleucher, Griesinger et al. 1994) :

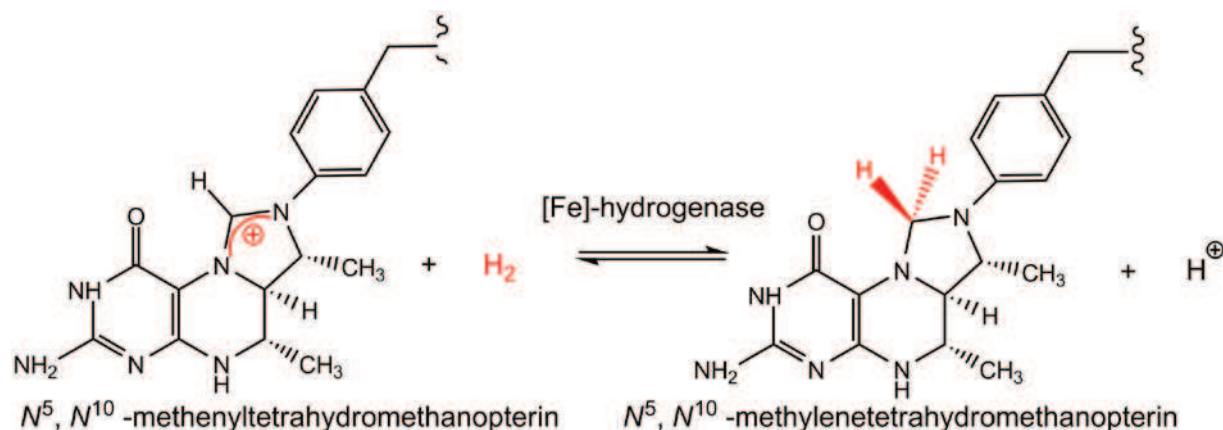


Figure I-1 R action catalys e par Hmd.

L'activit  de la prot ine est rendue possible gr ce   la pr sence d'un cofacteur unique, constitu  d'un atome de fer li    deux ligands CO, l'azote d'un d riv  de 2-pyridone, un groupement acyl li  au d riv  de 2-pyridone et un ligand inconnu (Figure I-2) (Hiromoto, Ataka et al. 2009).

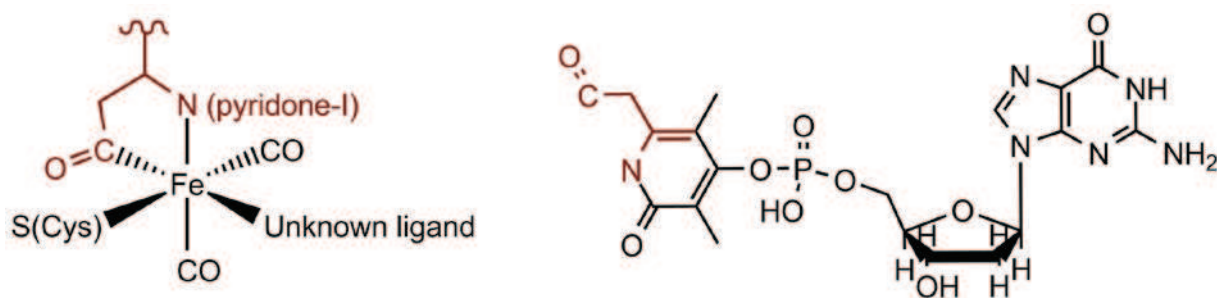


Figure I-2 Site actif d'Hmd.

Comme il l'a  t  discut  dans l'introduction, peu de travaux ont  t  r alis s sur l'assemblage de ce site organom tallique particulier. Cependant, par analogie avec les processus d'assemblage des centres FeS (Johnson, Dean et al. 2005), du FeMo-co de la nitrog nase (Rubio and Ludden 2008), des sites actifs des hydrog nases   fer-fer (Mulder, Shepard et al. 2011) et   nickel-fer (Fontecilla-Camps, Volbeda et al. 2007), il semble probable qu'une machinerie sp cifique existe pour ce site. Une analyse des g nomes des

organismes contenant la protéine Hmd nous a permis de mettre en évidence des séquences codant pour des protéines spécifiques de ces organismes, à proximité du gène *hmd*. Parmi ces gènes potentiels, nous avons identifié une séquence particulière, codant pour une protéine qui présente de fortes similarités de séquence avec les protéines BioB et HydE (Fontecilla-Camps, Amara et al. 2009), (McGlynn, Boyd et al. 2010). Cette protéine qui appartient à la famille des Protéines à Radical SAM (PRS) (Sofia, Chen et al. 2001), a été nommée HmdB. La présence d'une PRS dans le même opéron que le gène *hmd* (voir Introduction) nous rappelle les machineries de biosynthèse du « cluster H » des hydrogénases à fer-fer (présence des PRS HydE et HydG) (Mulder, Shepard et al. 2011) et du « FeMo-co » de la nitrogénase (présence de la PRS NifB) (Rubio and Ludden 2008). Nous avons donc émis l'hypothèse que la protéine HmdB est impliquée dans la biosynthèse du cofacteur de Hmd. Nous avons alors entrepris l'étude structurale de cette protéine. Pour cela, nous avons choisi de cloner puis purifier différents orthologues de la protéine, issus des organismes hyperthermophiles *Methanopyrus kandleri* et *Methanocaldococcus jannaschii*. En parallèle, nous avons aussi sélectionné la protéine HmdB issue de l'organisme mésophile *Methanococcus maripaludis*. Ce dernier a pour avantage le fait que la température de croissance est de 37°C, ce qui permet de réaliser des tests fonctionnels à des températures plus facilement accessibles au niveau expérimental. De plus, des outils de génétique et de biologie moléculaire ont été développés chez cet organisme (Walters and Chong 2009), ce qui rend potentiellement possible une étude fonctionnelle plus complète. En revanche, les protéines issues d'organismes hyperthermophiles sont réputées plus stables, donc plus faciles à purifier. Elles sont aussi, en général, plus compactes et plus rigides à température ambiante, donc plus faciles à cristalliser. Enfin, la surexpression de ces protéines dans un organisme mésophile, ici *E. coli*, permet de réaliser une étape de purification par chauffage de l'échantillon.

I.1) Clonage du gène, expression et purification de HmdB native de *Methanopyrus kandleri* (MkHmdB)

La protéine MkHmdB appartient à la famille des PRS. Elle possède donc un cofacteur [4Fe-4S] couplé à la S-adénosyl-L-méthionine (SAM) qui lui permet de générer l'espèce radicalaire 5'-désoxyadénosyl lors de l'initiation de la réaction radicalaire (Sofia, Chen et al. 2001), (Nicolet and Drennan 2004), (Frey, Hegeman et al. 2008). L'incorporation quantitative du centre [4Fe-4S] lors de la surexpression est impossible si la machinerie de biosynthèse des centres FeS n'est pas elle aussi, soit activée, soit surexprimée (Nakamura, Saeki et al. 1999). Nous disposons au laboratoire d'un plasmide permettant la surexpression de la machinerie ISC

d'*E. coli*. Il s'agit du plasmide pRSFIcsDuet-1 issu d'un plasmide pRSFDuet-1 de Novagen. Nous avons donc dû cloner le gène *HmdB* dans un plasmide compatible avec le plasmide pRSFDuet-1 pour permettre une co-expression dans *E. coli*. Le gène de la protéine HmdB de *Methanopyrus kandleri* a été cloné dans un plasmide pEtDuet-1 pour produire le plasmide pETMK0016Duet-1. Dans un premier temps et afin d'éviter tout problème d'expression dû à un usage alternatif des codons entre les organismes *M. kandleri* et *E. coli*, nous avons opté pour une souche d'expression d'*E. coli* BL21 (DE3) codons plus RIL. Les cellules ont été transformées concomitamment avec les plasmides pETMK0016Duet-1 et pRSFIscDuet-1, puis sélectionnées grâce aux deux antibiotiques kanamycine (résistance portée sur le plasmide pRSFDuet-1) et ampicilline (résistance portée sur le plasmide pETDuet-1). L'expression dans du milieu TB pendant 3h à 37°C était satisfaisante.

La protéine *MkHmdB* (locus primaire MK0016, GI 19886279) possède 356 acides aminés et un poids moléculaire de 39.3 kDa. Son point isoélectrique est de 5.9. Nous n'avons pas choisi d'ajouter d'étiquette poly-histidine à la protéine lors de la construction du plasmide d'expression, afin d'éviter un artefact éventuel tel que la liaison de l'étiquette polyhistidine sur le centre FeS, qui remplacerait alors un ligand « réel ». La protéine ne peut donc pas être purifiée à l'aide d'une colonne d'affinité contenant du nickel. Il a donc fallu tester différents types de résines pour permettre une purification efficace de notre protéine. Lors des premiers essais de purification, nous avons sélectionné une résine échangeuse d'anions (Q-Sepharose) suivi d'une colonne d'exclusion stérique (Superdex 200 16/60 prepgrade). Dans ce protocole l'échantillon était traité à l'enzyme désoxyribonucléase I (DNaseI) de pancréas bovin (Sigma,

DN25, 107K7013) lors de la lyse des cellules par ultrasons. Cependant, l'échantillon obtenu après la colonne Q-sepharose contenait encore des fragments d'ADN. Nous ajoutons alors une seconde fois de la DNase I. L'analyse par SDS-PAGE de l'échantillon protéique ainsi obtenu indiquait la présence de différents contaminants (Figure I-3). En effet, comme il est possible de le voir sur la Figure I-3, l'échantillon présente deux bandes majoritaires correspondant à des poids moléculaires de 40 et 35 kDa. La première bande a été attribuée à la protéine *MkHmdB* (poids moléculaire de 39.3 kDa). La seconde, d'intensité comparable, a été attribuée à un produit de dégradation de la protéine. Deux autres contaminants majeurs étaient présents avec des masses apparentes de 70 et 20 kDa.

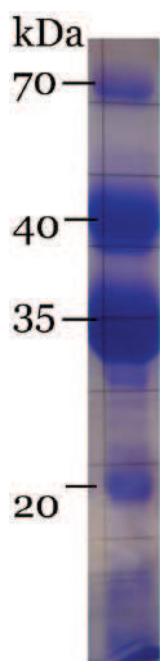


Figure I-3 Protéolyse de la protéine *MkHmdB*.

SDS-PAGE de l'échantillon contenant la protéine HmdB. Environ 10 µg de protéine déposés.

Lors de précédents essais de purification de la protéine HmdB de *M. maripaludis* (MpHmdB), il avait été observé que la protéine, purifiée dans des conditions similaires, avait été sujette à une protéolyse identique, malgré l'utilisation de « cocktails » d'inhibiteurs de protéases. Une analyse de séquence d'extrémité N-terminal de protéine à partir du fragment de 35 kDa, avait révélé que la position du site de coupure se situait entre les résidus Phe83 et Ala84 (triangle bleu sur la Figure I-4), contenus dans la séquence conservée « YCGFAAG » qui suit le motif Cx₅Cx₂C (Figure I-4) responsable de la fixation du centre [4Fe-4S] des PRS. Il semblait probable que la présence d'un tel motif Cx₅Cx₂C au lieu du motif canonique Cx₃Cx₂C des PRS, augmente la flexibilité de cette boucle et rend les protéines HmdB extrêmement sensibles à la protéolyse.

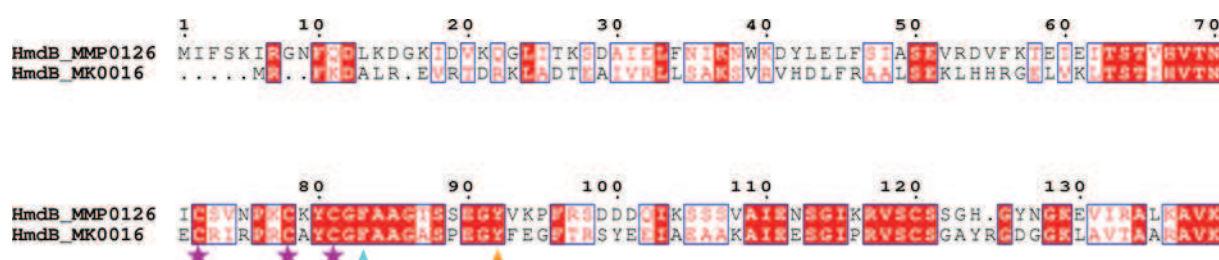


Figure I-4 Alignement de séquences de la protéine HmdB de *Methanococcus maripaludis* S2 et *Methanopyrus kandleri*.

HmdB de *Methanococcus maripaludis* S2 : HmdB_MMP0126 ; HmdB de *Methanopyrus kandleri* : HmdB_MK0016. Les triangles bleus et oranges indiquent respectivement le résidu précédant le site de protéolyse (obtenu par analyse de séquence N-terminal de protéine) de HmdB_MMP0126 et HmdB_MK0016. Les 3 cystéines du motif Cx₅Cx₂C sont indiquées par des étoiles violettes.

La stricte conservation de cette phénylalanine semble indiquer que cet acide aminé joue un rôle dans la fonction de la protéine. Cependant, afin de supprimer ce site de protéolyse, nous avons réalisé une mutation ponctuelle pour remplacer la phénylalanine par un acide aminé hydrophobe de taille comparable mais non aromatique : la leucine ou l'isoleucine. La purification de la protéine ainsi mutée nous a montré que le même type de protéolyse avait toujours lieu. Nous avons donc cherché à identifier ce site de coupure dans la protéine MkHmdB, puisque la boucle contenant le motif Cx₅Cx₂C présente d'autres acides aminés aromatiques, cibles potentiels de cette protéase.

Les protéines des différentes bandes du gel de la Figure I-3 ont été séquencées à leur extrémité N-terminale (voir Matériel et Méthodes). La bande à 20 kDa correspond à la chloramphenicol acétyltransférase (extrémité N-terminal identifiée : « MEKKIT »). La présence de ce contaminant est due à l'utilisation des cellules BL21 (DE3) codons plus RIL résistantes au chloramphénicol lors de l'expression de la protéine HmdB. Celle à 40 kDa correspond à la protéine HmdB (extrémité N-terminal identifiée : « MRFKDALR »), et l'extrémité N-terminal de la bande à 70 kDa n'a pas réussi à être identifiée. La séquence de l'extrémité N-terminal de la protéine à 35 kDa (« FEGFTRSY ») correspond, comme nous l'avions supposé, à un fragment de la protéine HmdB, coupé entre les résidus Tyr84 et Phe85 (triangle orange sur la Figure I-4).

La protéase responsable de la coupure des protéines *MpHmdB* et *MkHmdB* cible donc les résidus aromatiques. Il semblait évident que cette protéase corresponde à la chymotrypsine. En effet, cette dernière catalyse l'hydrolyse de la liaison peptidique qui suit un résidu aromatique F, Y et W, ainsi que la leucine (source <http://enzyme.expasy.org>). Nous avons alors suspecté notre lot de DNase I issue de pancréas de bœuf. Le pancréas est un organe sécréteur d'enzymes protéolytiques agissant dans le tube digestif. La chymotrypsine en fait partie. Nous avons alors remplacé la DNase I par de la Benzonase Nuclease (E1014, Sigma), qui est une endonucléase recombinante surexprimée dans *E. coli*. Le problème de contamination par des protéases provenant de pancréas est donc dans ce cas évité. Effectivement, lors des essais de purification suivants, la bande à 35 kDa n'était plus présente sur les gels. En revanche, malgré l'ajout d'étapes de purification, il n'était pas possible d'éliminer les autres contaminants de manière efficace. Nous avons donc choisi d'utiliser la souche d'*E. coli* BL21 (DE3) afin de ne pas surexprimer la chloramphenicol acétyltransférase. L'usage différentiel des codons n'a pas eu l'air d'affecter la surexpression de *MkHmdB* comme on peut le voir sur la Figure I-5. L'expression se fait de manière satisfaisante dans du milieu TB pendant 3h à 37°C avec l'apparition d'une bande à environ 40 kDa sur SDS-PAGE pour les échantillons après induction, ce qui correspond au poids de la protéine.

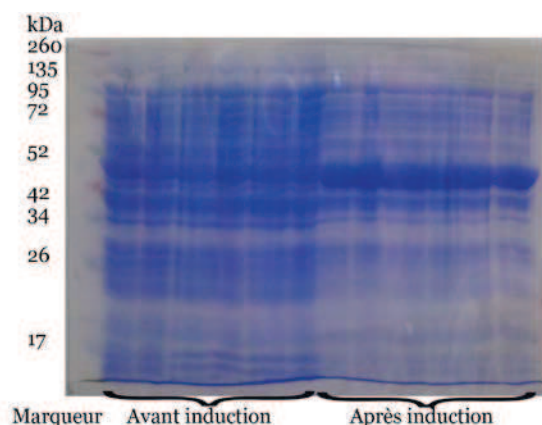


Figure I-5 Expression de la protéine *MkHmdB* native.

SDS-PAGE d'expression. De gauche à droite : le marqueur avec les poids moléculaires des protéines standard, 6 échantillons avant induction et 6 après induction.

Pour éliminer le plus de contaminants possibles, nous avons décidé d'utiliser une propriété potentielle de la protéine HmdB. En effet, cette dernière possède un fort pourcentage de similarités de séquences en acides aminés avec les protéines HydE et BioB, ce qui laisse penser que ces trois protéines ont un repliement identique (Nicolet and Drennan 2004). Or les protéines BioB et HydE ont la propriété d'interagir avec le bleu Cibacron contenu dans les résines Blue Sepharose. Nous avons donc testé cette résine pour purifier HmdB.

La première étape du protocole optimisé (voir Matériel et Méthodes) correspond à la lyse des cellules par ultrasons en présence de Benzonase nucléase. La seconde étape utilise la propriété de thermostabilité de *MkHmdB*, et correspond à une étape de chauffage de l'échantillon à 70°C pendant 20 min. A cette température les protéines contaminantes d'*E. coli*, organisme de surexpression de la protéine d'intérêt, sont dénaturées et peuvent être éliminées par centrifugation. La troisième étape fait appel à la résine Blue Sepharose. La protéine est éluée en un pic assez large dans le gradient.

La protéine ainsi obtenue est pure comme on peut le voir sur le gel de la Figure I-6, avec un rendement de 10 mg de protéine pour 2 L de culture, ce qui est satisfaisant pour réaliser les premiers essais de cristallisation.

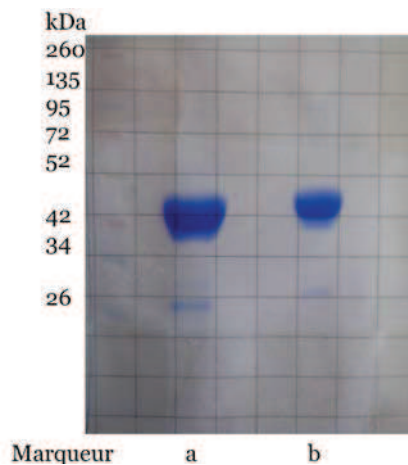


Figure I-6 Echantillon protéique final.

SDS-PAGE sur l'échantillon protéique final. De gauche à droite : le marqueur avec les poids moléculaires des protéines standard, a : 17.2 µg de protéine déposés, b : 8.6 µg déposés.

Le dosage en atomes de fer de l'échantillon indique la présence de 1.8 ± 0.2 atomes de fer par monomère de HmdB. Or nous attendons 4.0 atomes de fer par molécule de HmdB au minimum. Cependant la couleur marron de l'échantillon semble confirmer la présence de centres FeS (bandes de transfert de charge de liaison Fe-S). L'analyse par spectroscopie UV-visible (Figure I-7) montre un pic à 280 nm dû à la présence d'acides aminés aromatiques, ainsi qu'un épaulement à 320 nm et un pic large à 420 nm. Ces derniers sont caractéristiques de la présence de centres [4Fe-4S]. Ainsi, malgré la surexpression de l'opéron ISC d'*E. coli* et la purification de la protéine en condition anaérobie, l'incorporation du centre [4Fe-4S] n'a

pas été complète. Il est aussi possible que certaines molécules d'HmdB aient perdu leur centre [4Fe-4S] durant la purification, notamment lors de l'étape de chauffage hors de la boîte à gants, ceci malgré l'utilisation de tubes étanches à l'air. Afin d'avoir un échantillon homogène de la protéine, nous avons dû avoir recours à une étape de reconstitution *in vitro* du centre FeS (voir Matériel et Méthodes). L'échantillon après reconstitution contient alors 3.3 ± 0.4 atomes de fer par monomère de HmdB. L'analyse par spectroscopie UV-visible (Figure I-7) montre la présence des mêmes pics qu'avant reconstitution, caractéristiques de la présence de centres [4Fe-4S]. Le rapport DO à 280 nm / DO à 400 nm est plus faible, la reconstitution a donc augmenté la quantité de protéines contenant un centre [4Fe-4S].

L'analyse par comparaison des séquences de HmdB montre la présence de 7 cystéines conservées (une des cystéines l'est seulement parmi les organismes méthanogènes) (Figure I-17). Trois d'entre elles appartiennent au motif Cx_5Cx_2C responsable de la fixation du centre [4Fe-4S] « Radical SAM ». Les quatre autres pourraient être ligands d'un second centre du type [4Fe-4S] ou [2Fe-2S]. Cependant, le dosage en atomes de fer nous indique que nous n'arrivons pas à fixer plus de 4 atomes de fer par monomère de HmdB. Au cours de notre travail, une équipe concurrente a publié (en 2010) une caractérisation préliminaire de la protéine *MpHmdB* (McGlynn, Boyd et al. 2010). Ils obtiennent des résultats similaires quant au contenu en atomes de fer de la protéine après reconstitution *in vitro*. Il semble donc que les cystéines supplémentaires ne sont pas impliquées dans la fixation d'un autre centre FeS, du moins pas lors d'une reconstitution *in vitro*.

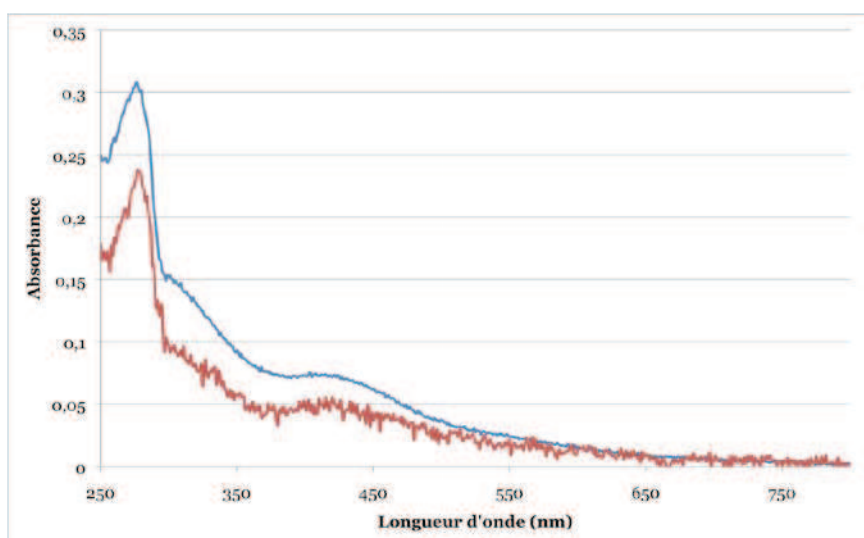


Figure I-7 HmdB avant et après reconstitution.

Spectres UV-visible de la protéine avant (en rouge) et après (en bleu) reconstitution. Les courbes ont été normalisées par rapport à la concentration en protéine déterminée par dosage au rose bengal. L'absorbance à 280 nm de la courbe bleue est plus importante car de la SAH a été ajoutée à la protéine.

I.2) Essais de cristallisation

Les essais de cristallisation de la protéine HmdB ont été menés avec la protéine à 5 mg/mL. En effet, il est apparu, lors des tests préliminaires, que cette concentration était la plus adaptée pour les conditions de cristallisation testées. Dans un premier temps, différentes matrices classiques de conditions de cristallisation (soit 200 conditions au total) ont été testées en condition anaérobie (utilisation de la BAG), afin de couvrir différents types de précipitants (polyéthylène glycols, sulfate d'ammonium, méthyl pentanediol etc.) et de pH. Ces premiers essais nous ont permis de déterminer les conditions les plus favorables à l'apparition de cristaux. Dans cette optique, un plan factoriel incomplet de 54 conditions a été réalisé en faisant varier les paramètres suivants :

- le choix du tampon : pH 6.5, 7.5, 8.5,
- le choix du précipitant : MPD, PEG 4000, sulfate d'ammonium,
- le choix possible d'un détergent : CHAPS, LDAO, n-Octyl- β -D-Glucopyranoside.

Les conditions les plus prometteuses (micro-cristaux, séparation de phase...) ont été à nouveau testées en faisant varier ces différents paramètres dans une gamme restreinte autour de la condition initiale. Cependant, malgré nos efforts, aucun cristal n'a été obtenu. Nous avons alors décidé d'utiliser le robot de cristallisation de l'institut pour couvrir rapidement une plus grande quantité de conditions différentes. Il est évident que cette approche ne permet pas de maintenir l'intégrité du centre [4Fe-4S] puisque ce dernier est sensible à l'oxygène. Notre souhait était de pouvoir obtenir des cristaux utilisables de la protéine, puis de transposer cette condition en condition anaérobie pour obtenir une forme de la protéine contenant le centre [4Fe-4S]. Nous avons ainsi pu tester plus de 800 conditions différentes. Cette approche s'est avérée fructueuse puisque nous avons pu obtenir des cristaux de protéine dans les conditions initiales de criblage (0.1 M citrate de sodium pH 5.6 ; 10% (w/v) PEG MME 2000, 3% (w/v) PEG 200, du screen MbClass Suite de Qiagen). Les cristaux ainsi obtenus sont incolores (Figure I-8), ce qui indique l'absence de centre FeS. Ceci était attendu, du fait de réaliser notre criblage en condition aérobie. Cependant, la présence de tampon citrate dans la condition de cristallisation est aussi problématique. En effet, le citrate est un chélateur de fer connu (Glusker 1980). Son utilisation, à des si fortes concentrations (100 mM) va donc favoriser la dégradation du centre [4Fe-4S] au cours du processus de cristallisation qui peut durer plusieurs semaines. Nous avons cependant essayé de reproduire manuellement ce résultat en condition anaérobie en BAG. Les mêmes cristaux incolores ont été reproduits. Le citrate joue donc son rôle de chélateur et déstabilise le centre [4Fe-4S], même en BAG. En revanche, la

substitution du citrate par une autre molécule jouant le même rôle tampon dans la même plage de pH ne permet pas d'obtenir de cristaux. Il semble donc que cette forme cristalline soit dépendante de l'absence de centre [4Fe-4S] sur la protéine. Par ailleurs, nos essais de reconstitution *in vitro* du centre [4Fe-4S] directement dans les cristaux n'ont pas été couronnés de succès. Récemment, un robot de cristallisation a été installé en BAG au laboratoire. Nous avons pu tester de nouveau un grand nombre de conditions de cristallisation différentes, mais malheureusement, seules les conditions contenant du citrate ont permis d'obtenir des cristaux toujours incolores. Il apparaît donc probable que la forme de la protéine avec son centre [4Fe-4S] ne soit pas compatible avec l'empilement cristallin obtenu. De plus, cette forme semble difficile à cristalliser, et nécessite plus de travail de caractérisation ainsi que de nouveaux essais de cristallisation. Ces points seront discutés plus tard.

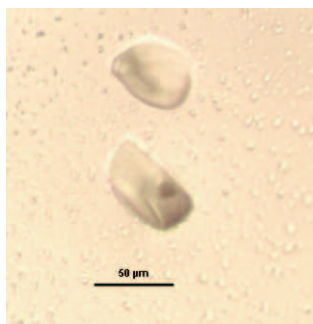


Figure I-8 Cristaux de HmdB.

I.3) Résolution de structure de HmdB

Le cristal décrit précédemment a été cryoprotégé par une solution 0.1 M cacodylate de sodium pH 5.6 ; 40% (v/v) PEG 400. Un jeu de données a été collecté à l'ESRF sur la ligne ID14-eh4. Les données ont été traitées avec le programme XDS (Kabsch 2010) (Tableau I–1).

Collecte et traitement de données		Affinement	
Ligne de lumière	ESRF ID14-eh4	R_{work}	0.232
Longueur d'onde (Å)	0.9395	R_{free}	0.292
Groupe d'espace	$P2_1$	Nombre d'atomes utilisés dans l'affinement (non hydrogène)	8789
Paramètres de maille (Å)	54.72 151.56 81.19	Molécules d'eau	94
	$\beta=100,8^\circ$	Ramachandran	
Résolution (Å)	2.53	régions les plus favorables (%)	91.0
hkl observés	132461 (18095)	régions permises supplémentaires (%)	7.4
hkl uniques	42735 (6712)	régions généreusement permises (%)	0.8
Redondance	3.1 (2.7)	régions interdites (%)	0.8
I/σ	9.76 (2.68)	Déviations "R.m.s."	
Complétude (%)	98.3 (96.2)	Liaisons (Å)	0.018
R_{sym} (%)	10.2 (40.8)	Angles ($^\circ$)	1.752

Tableau I–1 Statistiques de collecte, traitement de données et d'affinement pour *MkHmdB*.

Etant donné l'absence de centre FeS, il est impossible de résoudre la structure cristallographique de *MkHmdB* en utilisant la diffusion anormale (voir Matériel et Méthodes), sans avoir préalablement marqué la protéine avec par exemple des sélénométhionines. Une autre possibilité aurait été de diffuser des atomes lourds dans les cristaux pour utiliser la méthode du remplacement isomorphe (Taylor 2010). Cependant, comme HmdB présente de fortes similarités de séquence avec les protéines HydE et BioB pour lesquelles nous avons des structures cristallographiques disponibles (Nicolet, Rubach et al. 2008), (Berkovitch, Nicolet et al. 2004), nous avons opté pour la méthode du remplacement moléculaire (voir Matériel et méthodes). Le calcul du coefficient de Matthews (Matthews 1968) pour différents contenus de la maille laisse supposer que l'unité asymétrique contient 3 ($V_m=2.8 \text{ Da}/\text{\AA}^3$ soit 56% de solvant) ou 4 ($V_m=2.1$ soit 42% de solvant) molécules de HmdB.

Une superposition des structures de HydE (code PDB 3CIW, (Nicolet, Rubach et al. 2008) et BioB (code PDB 1R30, (Berkovitch, Nicolet et al. 2004) a permis de construire un premier modèle structural contenant uniquement les parties communes aux deux structures. Les séquences en acides aminés de ces deux modèles partiels ont été converties en chaîne poly-alanine, poly-sérine ou poly-glycine. Les premiers essais de remplacement moléculaire avec ces modèles ont été infructueux et ont révélé les limites de cette approche dans notre cas particulier. En effet, le groupe d'espace $P2_1$ possède une origine indéfinie suivant l'axe **b**. La première solution (translation) n'est donc, pas autant contrastée. Par ailleurs, la structure en tonneau $(\beta\alpha)_8$ présente des éléments de pseudo symétrie d'ordre 2, 4 et 8, ce qui produit des solutions fausses. La technique du remplacement moléculaire est très sensible. Ainsi, une hélice α dont l'orientation serait légèrement différente diminue fortement le score d'une solution pourtant correcte. Nous avons alors entrepris l'élaboration d'un nouveau modèle. Nous avons donc, dans un premier temps, aligné la séquence de *MkHmdB* sur l'alignement de séquence réalisé à partir des structures de HydE et BioB (Nicolet, Rubach et al. 2008). Cet alignement de séquence, ainsi que les structures cristallographiques de HydE et BioB, ont été utilisés pour créer un modèle moyen à l'aide du programme MODELLER (Fiser and Sali 2003). Ce modèle a alors été converti en chaîne poly-alanine avant d'effectuer un nouveau cycle de recherche de solutions par remplacement moléculaire. En parallèle, nous avons utilisé le programme POLARRFN (<http://www.ccp4.ac.uk/html/polarrfn.html>) pour détecter les possibles axes de symétrie non cristallographique. Trois pics significatifs dans la fonction de self Patterson ont ainsi pu être mis en évidence. Il s'agit des axes suivants :

- | | |
|--|---|
| - $(\omega, \phi, \kappa) = (90, 90, 180)$ | intensité normalisée : 100% du pic à l'origine |
| - $(\omega, \phi, \kappa) = (1.6, 0, 180)$ | intensité normalisée : 74.9% du pic à l'origine |
| - $(\omega, \phi, \kappa) = (90, 0, 3.1)$ | intensité normalisée : 73.8% du pic à l'origine |

Ces valeurs indiquent la présence de deux axes de symétrie non cristallographique d'ordre 2 perpendiculaires à l'axe 2_1 cristallographique, ainsi qu'une pseudo translation. Ce résultat indique aussi que le contenu de l'unité asymétrique est constitué de 4 molécules d'HmdB et non pas de 3. De plus, nous avons pu sélectionner les solutions de remplacement moléculaire proposées par le programme PHASER (McCoy, Grosse-Kunstleve et al. 2007) pour vérifier qu'elles étaient compatibles avec ces relations de symétrie non cristallographiques. Tout ceci nous a permis d'obtenir une première solution crédible, même si elle était rejetée par le programme pour cause de mauvais contacts cristallins. Nous avons alors pu calculer une première carte de densité électronique. Il est apparu que l'hélice α_8 était significativement

décalée dans notre modèle par rapport à la densité électronique. Après correction manuelle de la position et de l'orientation de cette hélice dans notre modèle, nous avons réitéré notre recherche par remplacement moléculaire. Cette fois, une solution comprenant 4 molécules est apparue clairement contrastée. Elle respectait les symétries non cristallographiques précédemment déterminées. Cependant, la qualité du modèle initial étant tout de même insuffisante, les cartes de densité électronique obtenues étaient de mauvaise qualité et difficilement interprétables pour une reconstruction manuelle. Nous avons alors échaffaudé une procédure itérative de modification de densité électronique en incluant l'aplatissement de solvant (voir Matériel et Méthodes) et la moyennation sur les 4 molécules de l'unité asymétrique (voir Matériel et Méthodes), à l'aide du programme DM (ftp://ftp.ccp4.ac.uk/ccp4/newsletter/nov_94/09_cowtan.ps). Les cartes de densité électronique ainsi obtenues permettaient de corriger petit à petit le modèle initial. Ce dernier était alors réutilisé pour effectuer un remplacement moléculaire, afin d'améliorer la précision des opérateurs de symétrie non cristallographique, eux même réutilisés pour effectuer une nouvelle étape de modification de densité électronique. La qualité des cartes obtenues permettait de progresser dans la reconstruction du modèle. Lorsque la qualité du modèle a été suffisante, nous avons pu abandonner cette procédure au profit de la procédure classique d'affinement avec le programme REFMAC5 (Murshudov, Vagin et al. 1997), suivi d'étapes de reconstruction manuelle. Les étapes de la résolution de structure sont résumées dans le modèle ci-dessous (Figure I-9).

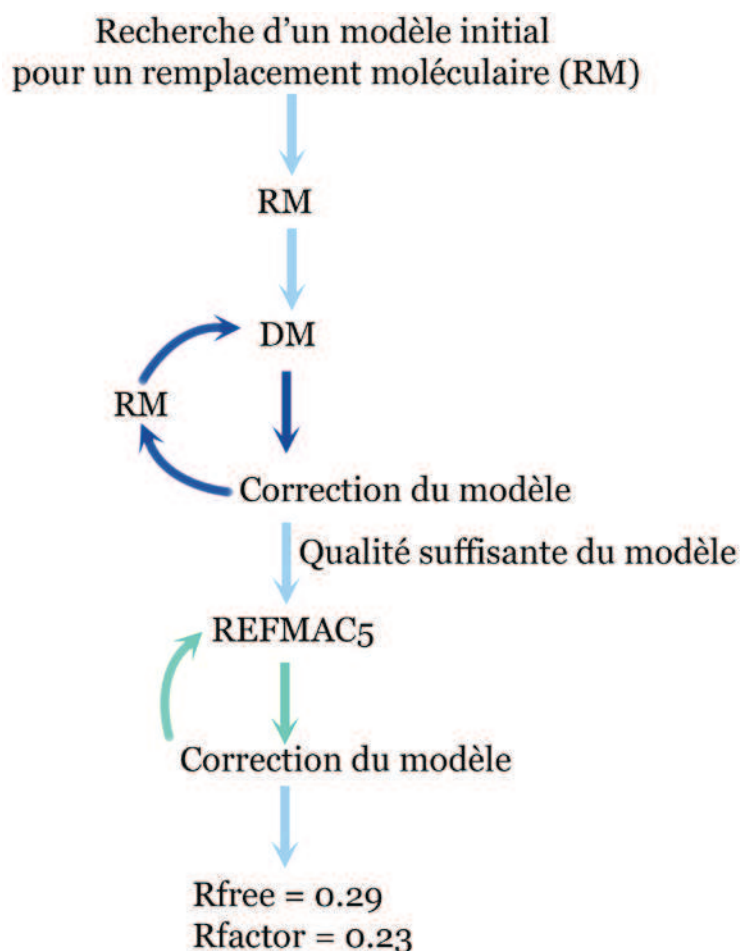


Figure I–9 Résolution de la structure de *MkHmdB*.

I.4) Analyse structurale de la protéine HmdB

I.4.1) Résidus non visibles dans sa structure

Les cristaux de la protéine HmdB comportent 4 monomères par unité asymétrique. Plusieurs groupes de résidus ne sont pas visibles dans sa structure :

- les résidus 64 à 87, dont font partie les 3 cystéines fixant le centre [4Fe-4S],
- les résidus 240 à 248, 303 à 315, et 336 à l'extrémité C-terminal.

Il n'y pas de centre [4Fe-4S] dans la structure, ce que nous supposons d'après la couleur des cristaux. Malgré son ajout avant cristallisation, la SAH (analogue de la SAM) n'est pas présente.

La comparaison de la structure de *MkHmdB* avec celles de HydE (code PDB 3iiz, (Nicolet, Amara et al. 2009) et BioB (code PDB 1r30, (Berkovitch, Nicolet et al. 2004) montre que les résidus non visibles se situent dans des boucles sur le haut du tonneau (Figure I–10).

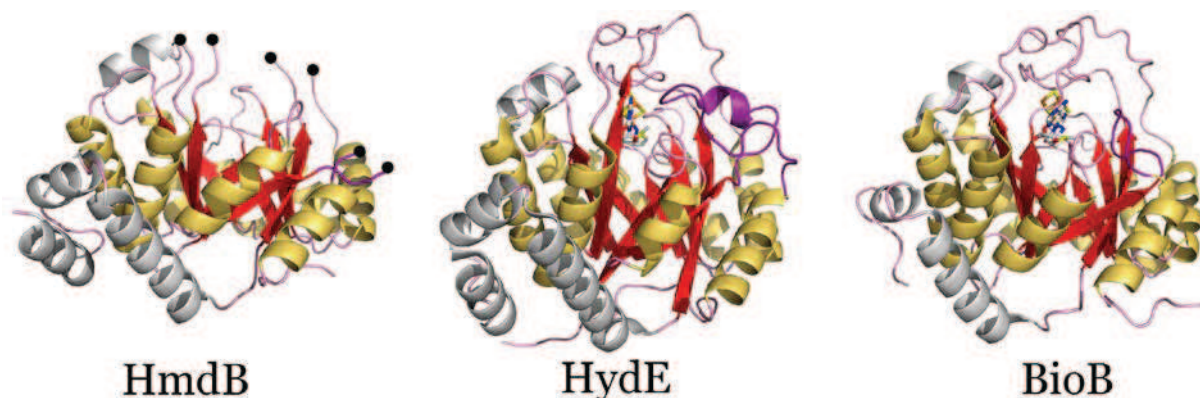


Figure I-10 Structures des protéines à radical SAM ($\beta\alpha$)₈.

Les brins β et hélices α composant le tonneau triosephosphate isomérase sont représentés en rouge et jaune respectivement. Les hélices α supplémentaires sont en gris clair et les boucles en rose. La boucle fermant une extrémité du tonneau (« couvercle ») est représentée en violet. De gauche à droite : structure de HmdB avec les résidus aux extrémités des boucles non visibles représentées par un cercle noir ; HydE et BioB avec leur centre [4Fe-4S] et SAM représentés en bâtons.

L'alignement structural des protéines HmdB, HydE et BioB (Figure I-11) montre que les résidus non visibles dans la structure se situent dans des boucles localisées entre les hélices et brins du tonneau. La boucle entre le brin β 8 et l'hélice α 8, que l'on appelle « couvercle », couvre une extrémité du tonneau dans les structures de BioB et HydE (en violet sur la Figure I-10).

Dans HydE le « couvercle » possède une séquence conservée YxxY. Le rôle de ces tyrosines a été étudié par mutagenèse dirigée (Nicolet, Rubach et al. 2008). Comme la fonction de HydE n'est pas connue, l'effet de la mutagenèse a été étudié *in vivo* en mesurant l'activité de l'hydrogénase recombinante de *C. acetobutylicum* dans *E. coli*. La coexpression hétérologue des protéines HydE, HydF, HydG et HydA de *C. acetobutylicum* dans *E. coli* amène à l'obtention d'hydrogénases à Fe-Fe actives. Dans le cas d'une mutation affectant l'activité de HydE, l'activité hydrogénase devrait donc diminuer voire être inexistante. Les auteurs ont observé que la mutation des deux résidus tyrosine en phénylalanine n'affecte pas l'activité hydrogénase, tandis que la mutation en alanine la supprime. Les résidus aromatiques sont peut-être essentiels pour protéger le site actif du solvant lors de la catalyse.

Dans BioB le « couvercle » possède deux thréonines (Thr292 et 293 dans *E. coli*) qui font partie des résidus interagissant avec son substrat, la déthiobiotine (DTB) (Berkovitch, Nicolet et al. 2004). La boucle reposerait donc sur le tonneau après la liaison de la DTB au site actif, le protégeant du solvant.

Figure I-11 Alignement structural des protéines *MkHmdB*, *EcBioB*, *TmHydE*.



Alignement structural des protéines *MkHmdB*, *EcBioB* (code PDB 1r30) et *TmHydE* (code PDB 3iiz) réalisé avec CLUSTALW2 (Larkin, Blackshields et al. 2007), corrigé manuellement pour tenir compte de la localisation structurale des résidus et de leur conservation spatiale, et représenté avec le logiciel ESPript (Gouet, Courcelle et al. 1999). Les boîtes en fond rouge et fond blanc indiquent respectivement les acides aminés conservés et similaires. Les séquences caractéristiques aux protéines à radical SAM sont indiquées par un trait rouge. Les traits noirs indiquent les résidus de HmdB non visibles dans sa structure.

I.4.1) Les limites du modèle de la protéine HmdB

La structure de la protéine HmdB a été résolue sous sa forme apo et des boucles du haut du tonneau sont absentes. Lorsque l’on superpose la structure d’HmdB sur celles de BioB et HydE, on peut définir une partie homologue entre elles au bas du tonneau. La partie homologue d’HmdB est représentée en blanc sur la Figure I–12. La valeur de rmsd (root mean square deviation) de la superposition de cette partie sur BioB et HydE est plus faible que celle de la protéine entière. Toutes les valeurs sont indiquées dans le Tableau I–2.

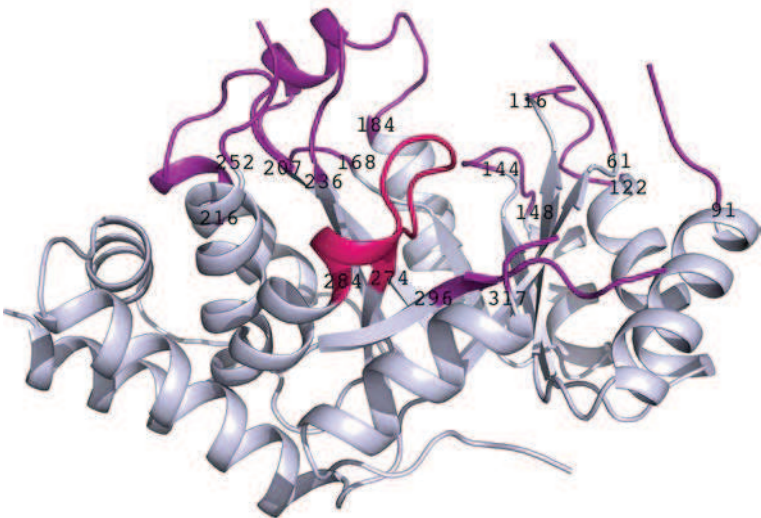


Figure I–12 Limites du modèle de la protéine HmdB.

En blanc, la partie homologue à BioB et HydE. En violet, les résidus dont la conformation serait identique dans la structure de l’holoprotéine. En rose, les résidus dont la conformation serait différente dans la structure de l’holoprotéine.

rmsd (Å)	BioB	HydE
HmdB	1.83	1.98
HmdB (partie homologue uniquement)	1.65	1.72

Tableau I–2 Superposition de la protéine HmdB sur BioB ou HydE.

Valeurs calculées avec le logiciel SSM (Krissinel and Henrick 2004).

Du fait de l'absence du centre [4Fe-4S], de la SAH et de la non-visibilité de résidus contenus dans des boucles sur le haut du tonneau, il est nécessaire d'analyser la partie intermédiaire de la structure représentée par les résidus en violet et rose sur la Figure I-12. Nous avons comparé la conformation de ces résidus par rapport à celles des protéines HydE et BioB (Tableau I-3), dans le but définir les résidus dont la conformation serait identique dans la structure de l'holo-HmdB.

HmdB	HydE	BioB	Remarques
(61 à 63)- (88 à 91)	60 à 85	49 à 77	Il s'agit de la boucle contenant les 3 cystéines conservées qui lient le centre [4Fe-4S]. Elles ont une conformation similaire dans HydE et BioB. La conformation des résidus visibles de HmdB est similaire à celles de BioB et HydE.
116 à 122	110 à 117	102 à 111	La conformation des résidus de cette boucle est différente dans HydE et BioB. La position de la boucle d'HmdB se trouve entre celles de BioB et HydE.
144 à 148	138 à 141	132 à 135	Conformations similaires.
168 à 184	161 à 177	155 à 170	Conformations similaires. De plus, l'hélice α 4a est aussi présente dans HmdB.
207-216	200 à 210	193 à 202	Conformations similaires.
(236 à 239)- (249 à 252)	229 à 245	223 à 239	La conformation de la boucle dans HydE est similaire à celle dans BioB. La conformation des résidus visibles de HmdB est similaire à celles de HydE et BioB.
274 à 284	267 à 279	261 à 273	Voir le texte principal.
(296 à 302)- (316 à 317)	291 à 319	285 à 297	Voir le texte principal.

Tableau I-3 Comparaison de la conformation de résidus situés sur une partie intermédiaire entre le haut et le bas du tonneau des structures de HmdB, BioB et HydE.

Afin de définir les résidus dont la conformation serait identique à une structure de l'holoprotéine. Les numéros de résidus analysés sont donnés.

Dans la boucle 296 à 317 (291 à 319 pour HydE, 285 à 297 pour BioB), seuls les résidus 296 à 302 et 316 à 317 sont visibles. Il s'agit du « couvercle » fermant un côté du tonneau. Le nombre de résidus dans celui-ci est variable, et sa conformation est différente dans HydE ou BioB. La conformation des résidus visibles de HmdB est aussi différente. En ce qui concerne la boucle 274 à 284 (en rose sur la Figure I-12) (267 à 279 pour HydE, 261 à 273 pour BioB), elle a de mauvais contacts avec les résidus du « couvercle » de HydE et BioB. Si ce dernier

dans HmdB adoptait une conformation identique à celle de BioB ou HydE dans l'holoprotéine, il y aurait donc une mauvaise interaction. Les résidus présents dans le « couvercle » n'étant pas similaires et en nombre différent dans les trois protéines, il est difficile de savoir si le couvercle d'HmdB adopterait la même conformation que celle d'HydE ou BioB.

La boucle 274 à 284 adopte une conformation très différente de celle de BioB et HydE. Le désordre de la boucle 296 à 302 semble donc être responsable de la position de cette boucle. Il semble y avoir eu un réarrangement du fait de l'absence du centre [4Fe-4S].

D'après l'analyse faite, seul les résidus 274 à 284 semblent adopter une conformation différente de celle qu'ils adopteraient dans l'holoprotéine.

1.4.2) Agencement spatial des motifs conservés des protéines à radical SAM dans MkHmdB

Ayant vu précédemment que le modèle de l'apoprotéine semble proche du modèle de l'holoprotéine dans ses parties visibles, mis à part pour les résidus 274 à 284, nous avons fait une comparaison de sa structure avec celles de BioB (code PDB 1r30) et HydE (code PDB 3iiz). La SAM a été ajoutée fictivement dans HmdB, en superposant sa structure avec celle de HydE, afin de pouvoir comparer plus facilement les trois structures. Les séquences protéiques mentionnées peuvent être retrouvées dans les Figure I-11 et Figure I-17.

Il existe plusieurs motifs conservés dans les protéines à radical SAM, comme déjà mentionné dans l'introduction. Le premier motif est la séquence Cx_3Cx_2C (Sofia, Chen et al. 2001), dédié à la liaison au centre [4Fe-4S]. Il se trouve dans la boucle entre le brin $\beta 1$ l'hélice $\alpha 1$. Il n'est pas visible dans notre structure, mais l'est dans les structures des protéines BioB et HydE (Figure I-13).

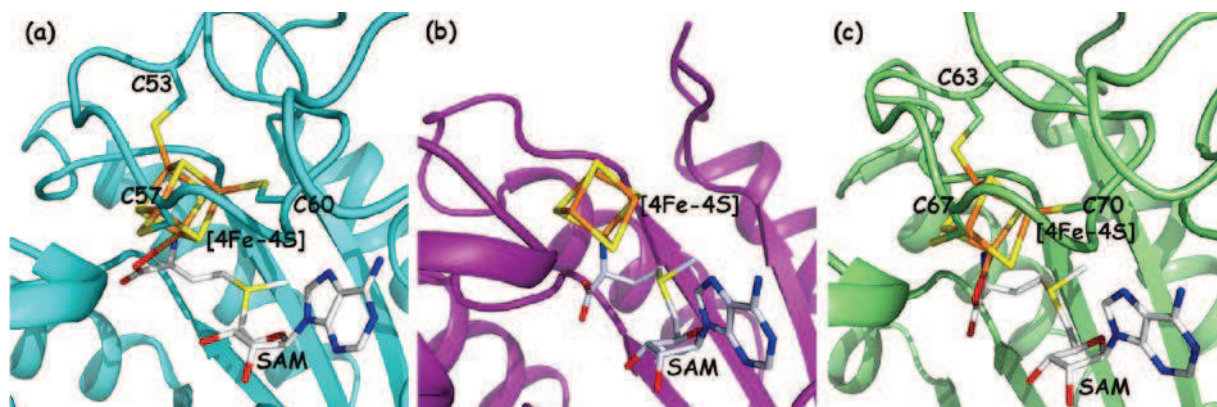


Figure I-13 Motif de fixation du centre [4Fe-4S].

Motif de fixation du centre [4Fe-4S] des protéines : (a) BioB, (b) HmdB, (c) HydE.

D'autres motifs conserv s ont  t  identifi s par Nicolet et Drennan en 2004 (Nicolet and Drennan 2004).

Le deuxi me motif « GxIxGxxE » est un motif riche en glycine (Figure I-14). Il se trouve dans le brin  5. Une interaction entre le r sidu hydrophobe « I » et l'ad nine de la SAM a  t  observ e par les auteurs. Les r sids Ile192 de BioB et Met199 de HydE correspondants, interagissent avec l'ad nine de la SAM. Le r sidu « I » de HmdB, la Leu106, poss de la m me orientation que celle des r sids « I » de HydE et BioB. Il semble donc qu'elle aurait le m me r le. Les r sids « G » (celui en 3 me position) et « E » sont dans la boucle suivant le brin  5. Ils semblent avoir un r le dans la structure de la boucle par la liaison entre l'azote du r sidu « G » et la chaine lat rale du r sidu « E », qui a l'air de contribuer au maintien du site de liaison   l'AdoMet. Cette liaison est observ e dans BioB (Glu197 et Gly194) et HydE (Gln205 et Gly201). Elle l'est aussi dans HmdB, entre la Glu211 et la G208. La boucle a d'ailleurs une conformation  quivalente   celle de BioB. Dans HydE un r sidu suppl mentaire non conserv  est ins r  dans cette boucle, juste avant la Gln205. La boucle a alors une conformation diff rente de celle de BioB et HmdB. D'autres interactions entre les r sids du motif peuvent avoir lieu, comme observ  dans d'autres structures de PRS (Vey and Drennan 2011).

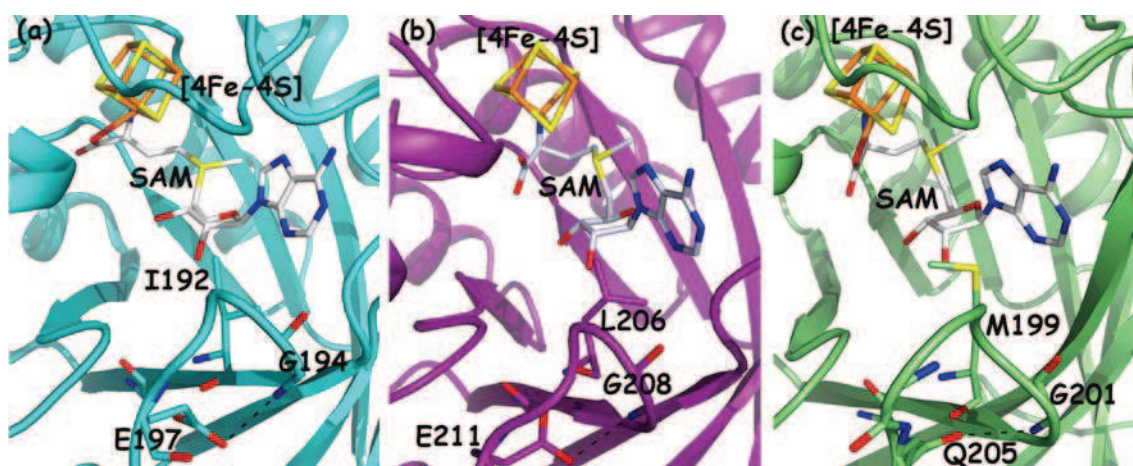


Figure I-14 Motif « GxIxGxxE ».

Motif « GxIxGxxE » des prot ines : (a) BioB, (b) HmdB, (c) HydE.

Le troisième motif « GGE » (Figure I-15) est riche en glycine, il se situe à la fin du brin $\beta 2$. Il apparaît important pour la conformation de la boucle suivant le brin $\beta 2$ pour permettre la liaison hydrogène avec l'azote du motif méthionine de la SAM. Dans BioB, les carbonyles des Ala100 et Trp102 le lient. Dans HydE, il s'agit du carbonyle de la Ser108. Dans HmdB, le carbonyle de la Gly114 est à la même position que celui des Ala100 de BioB et Ser108 de HydE, il semble qu'il aurait donc le même rôle.

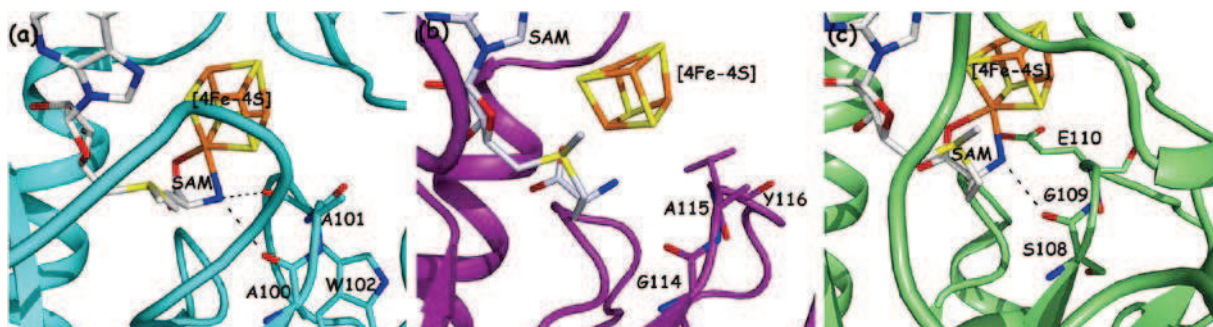


Figure I-15 Motif « GGE ».

Motif « GGE » des protéines : (a) BioB, (b) HmdB, (c) HydE.

Les motifs précédents sont ceux conservés dans les PRS, ayant pour conséquence une position similaire de l'adénine et l'amine du fragment méthionine de la SAM. Voici ce qui concerne la partie variable des résidus impliqués dans la liaison au carboxylate et la liaison au ribose. Un résidu arginine conservé lie et oriente le motif carboxylate de la SAM dans HydE et BioB. Il s'agit des Arg180 et Arg173, respectivement. Ce résidu se retrouve dans HmdB (Arg187), et est dans la même position que dans HydE et BioB, il pourrait donc orienter le groupement carboxylate de la SAM (Figure I-16). Cette arginine se trouve à proximité, dans l'espace, des motifs YN¹⁵¹HN¹⁵³LD¹⁵⁵T de BioB, L¹⁵⁷LR¹⁵⁹HE¹⁶¹T de HydE et C¹⁶⁴CN¹⁶⁶LE¹⁶⁸T(ou S) de HmdB (motif « CCNLET(ou S)/R ») (Figure I-16). Les résidus Asp155 de BioB et Glu161 de HydE lient le ribose de la SAM par une liaison hydrogène. La Glu168 de HmdB est à une position similaire. Dans BioB, l'Asn153 lie le ribose de la SAM et son substrat la déthiobiotine (non représentée dans la figure) ; l'Asn151 lie aussi le substrat. Le motif YNHNLDLT de BioB s'est révélé important dans sa fonction puisque la mutation en alanine des résidus N, H, N ou D élimine ou diminue fortement la production de biotine par la protéine (Lotierzo, Raux et al. 2006). Le rôle des résidus Asn153 et Asp155 a été étudié plus en détail (Farrar and Jarrett 2009) : leur mutation en alanine n'affecte presque pas la liaison de

déthiobiotine et la SAM ; c'est la production de biotine et 5'-désoxyadénosine qui est presque inexistante. La mutation de ces résidus en d'autres acides aminés rend quand même possible la production de 5'-désoxyadénosine, mais la production de biotine est toujours presque inexistante. Il semble que ces résidus aient un rôle dans le contrôle de la position et la trajectoire du radical 5'-désoxyadénosyl. Dans HmdB, la position de l'Asn166 est similaire à celle de l'Asn153, elle pourrait donc avoir les mêmes rôles. Dans HydE, la mutation de l'Arg159 en acide glutamique abolit son rôle de maturase de l'hydrogénase à [Fe-Fe] (Nicolet, Rubach et al. 2008). Il apparaît donc bien que ces résidus soient importants dans la catalyse et que les variations dans l'interaction à la SAM permettent à l'enzyme de spécifier sa réaction.

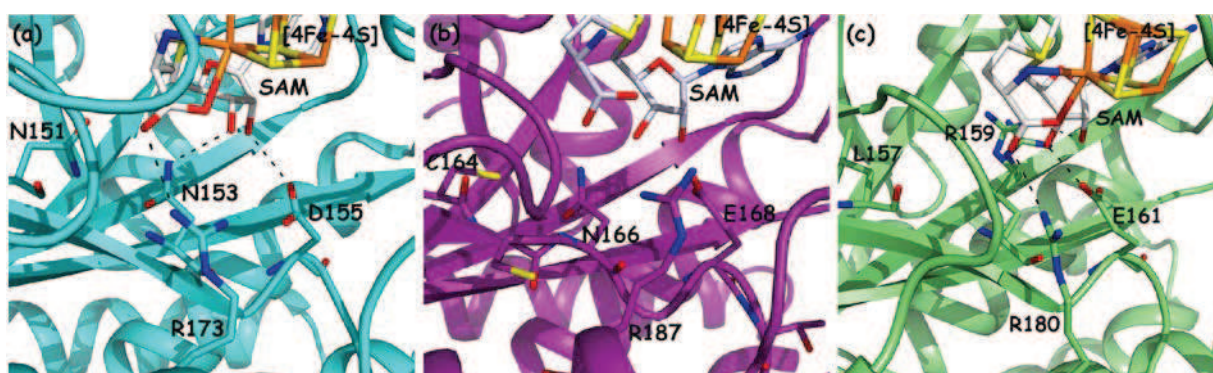


Figure I-16 Motif « CCNLET(ou S)/R » de HmdB, YNHNLDT/R de BioB et LLRHET/R de HydE.

(a) BioB, (b) HmdB, (c) HydE.

Methanocaldococcus_jannaschii	52	R	F	K	I	D	N	W	R	D	Y	L	K	F	D	I	S	R	C	D	Y	F	K	.	K	K		H	N	N	L	C	C		K	K	E	P	R	L																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		
Methanocaldococcus_vulcanius	36	K	R	F	K	I	D	N	W	R	D	Y	L	K	F	D	I	S	R	C	D	Y	F	K	.	R	E		H	N	N	L	C	C		K	K	E	P	R	L																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
Methanocaldococcus_sp.FS406-22	33	K	R	F	K	I	D	N	W	G	D	Y	L	K	F	N	I	S	R	C	D	Y	F	K	.	K	E		H	N	N	L	C	C		K	K	E	X	P	R	L																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																
Methanocaldococcus_fervens	33	K	R	F	K	I	D	K	W	D	Y	L	Q	.	F	R	L	S	R	C	D	Y	F	K	.	K	E		H	N	N	L	C	C		K	K	E	X	P	R	L																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																
Methanoterris_igneus_Kol5	33	K	R	F	K	I	N	W	S	D	Y	L	R	.	F	N	I	S	R	C	D	Y	F	K	.	K	E		H	N	N	L	C	C		K	K	E	X	P	R	L																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																
Methanothermococcus_okinawensis	32	N	T	F	I	A	H	N	W	D	Y	L	K	.	F	S	I	S	R	C	D	Y	F	K	.	K	E		K	N	N	N	L	C	C		K	K	E	T	P	R	L																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																															
Methanococcus_aeolicus_Nankai-	34	K	R	F	I	S	K	G	N	D	Y	L	N	.	F	S	I	S	R	C	D	Y	F	K	.	K	E		K	N	N	N	L	C	C		K	K	E	T	P	R	L																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																															
Methanococcus_maripaludis_C6	32	E	T	G	I	K	N	W	D	Y	L	E	.	F	S	I	S	R	C	D	Y	F	K	.	N	E		S	T	T	T	T	C	C		K	K	E	X	P	R	L																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																
Methanococcus_maripaludis_S2	32	E	T	F	I	K	N	W	D	Y	L	E	.	F	S	I	S	R	C	D	Y	F	K	.	T	E		S	T	T	T	T	C	C		K	K	E	X	P	R	L																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																
Methanococcus_vannielii	32	A	T	F	I	R	T	W	D	E	C	L	K	.	F	S	I	S	R	C	D	Y	F	K	.	P	E		S	T	T	T	T	C	C		K	K	E	X	P	R	L																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																															
Methanococcus_voltae	71	E	T	F	I	S	K	I	A	D	Y	L	E	.	F	S	I	S	R	C	D	Y	F	K	.	D			S	V	S	V	C	C		K	K	E	X	P	R	L																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																
Methanocaldococcus_infernus	27	K	R	F	I	S	D	K	V	R	D	Y	L	E	.	F	K	I	S	L	R	E	K	F	K	.	E	K		K	S	S	V	C	C		K	K	E	X	P	R	L																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																															
Methanocorpusculum_labreanum	22	N	T	F	S	S	S	S	Y	A	E	L	E	P	.	I	N	V	D	.	Q	K	S	G	H	.	S		H	T	T	T	C	C		K	K	E	T	P	R	L																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																
Methanoplanus_petrolearius	22	Q	T	F	S	N	P	A	D	E	E	F	A	R	.	I	E	T	F	N	S	R	D	N	.	K	F		H	T	T	T	C	C		K	K	E	T	P	R	L																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																
Methanothermobacter_t_str.DH	22	Q	T	F	I	R	E	S	G	D	L	A	S	.	M	K	T	F	N	M	R	E	N	.	G	S		O	R	R	G	G		K	K	E	N	S	F	K																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		
Methanothermobacter_m.	21	Q	T	F	I	Q	K	D	P	K	D	F	A	L	.	M	E	T	F	S	M	K	E	R	.	G	V		O	R	R	G	G		K	K	E	N	S	F	K																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
Methanothermobacter_t_str.W	22	E	T	F	I	R	N	R	P	E	E	F	A	D	.	M	D	T	F	N	M	D	H	.	D	R		O	R	R	T	O		K	K	E	T	P	R	L																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		
Methanobrevibacter_smithii	22	Q	T	F	L	K	I	N	K	E	E	D	L	K	.	C	E	T	V	S	R	N	K	H	.	S	.		S	T	T	T	T	C	C		K	K	E	X	P	R	L																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																															
Methanobrevibacter_ruminantium	22	Q	T	F	I	D	E	E	T	L	K	E	.	C	K	I	C	D	.	R	N	E	C	.	S			O	R	R	K	K		K	K	E	L	A	Y	K																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		
Methanothermus_fervidus	23	K	R	F	I	K	S	N	D	F	K	E	.	M	Y	T	S	.	R	D	R	Y	K	.	D	S		S	T	T	T	T	C	C		K	K	E	X	P	R	L																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																
Methanopyrus_kandleri	23	V	L	L	S	A	K	S	V	R	V	H	D	.	F	R	A	.	L	S	.	K	L	H	R	G	E		S	T	T	T	T	C	C		K	K	E	X	P	R	L																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																															
Desulfurobacterium_thermolitho	1

[illegible]

114

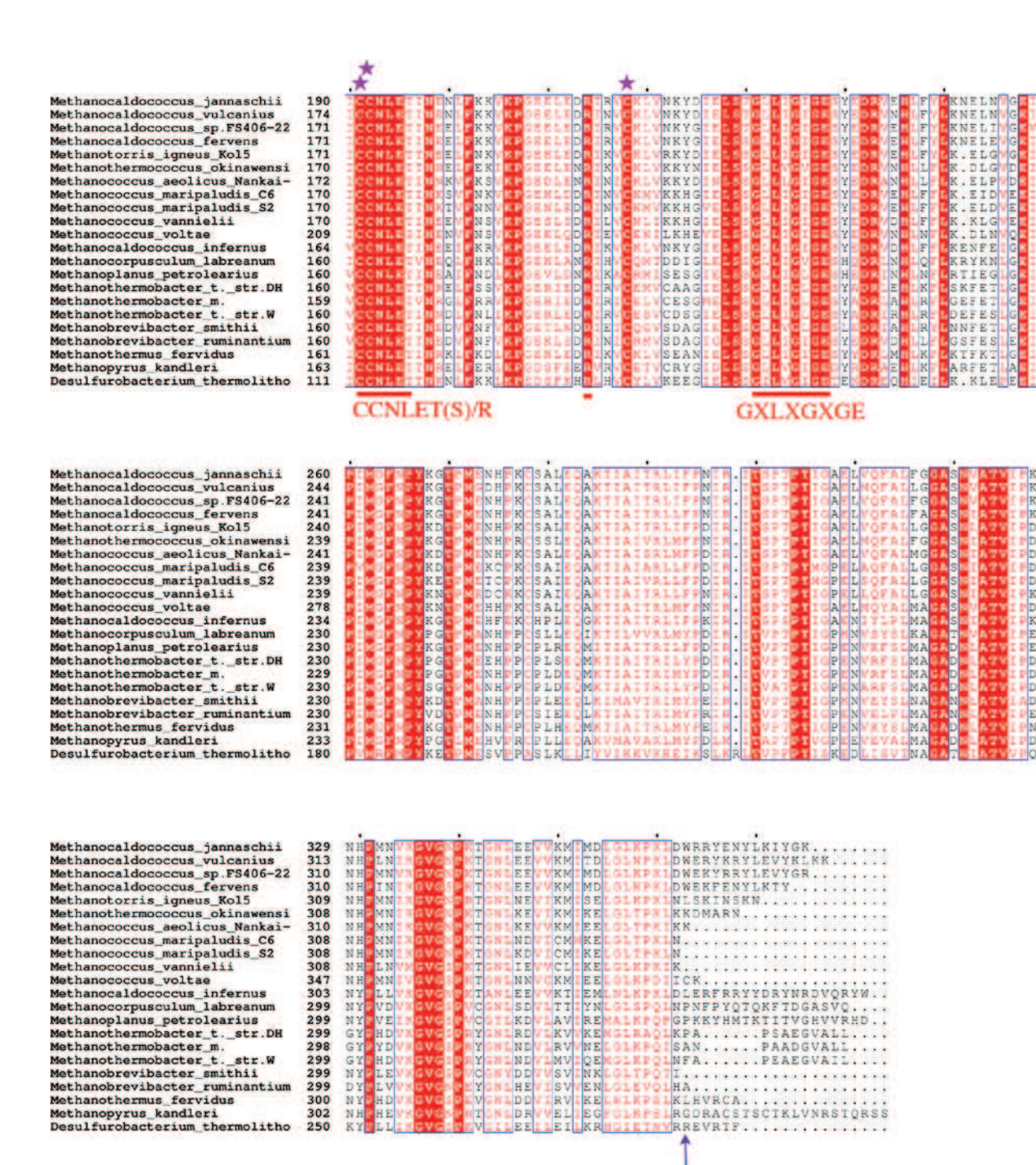


Figure I-17 Alignement multiple de séquences de protéines HmdB.

Alignement réalisé avec CLUSTALW2 (Larkin, Blackshields et al. 2007). Les boîtes en fond rouge et fond blanc indiquent respectivement les acides aminés conservés et similaires. Les séquences caractéristiques aux protéines à radical SAM sont indiquées par un trait rouge. Les étoiles violettes indiquent les résidus cystéine conservés. La flèche violette indique la G336.

Tous les motifs conservés des protéines à radical SAM, mis à part le motif de fixation du centre [4Fe-4S], sont bien positionnés structuralement dans *MkHmdB*. Le bas du site actif de la SAM semble donc être bien structuré dans HmdB malgré la flexibilité des boucles du dessus du tonneau. L'affinité de la SAM pour ce site ne semble pas assez importante

puisqu'elle n'est pas présente dans la structure. La liaison au centre [4Fe-4S] semble essentielle à sa présence.

I.5) Empilement cristallin de *MkHmdB*

Les cristaux de la protéine HmdB comportent 4 molécules par unité asymétrique. Plusieurs symétriques de ces 4 molécules ont été représentés dans la Figure I-18, définissant l'empilement cristallin.

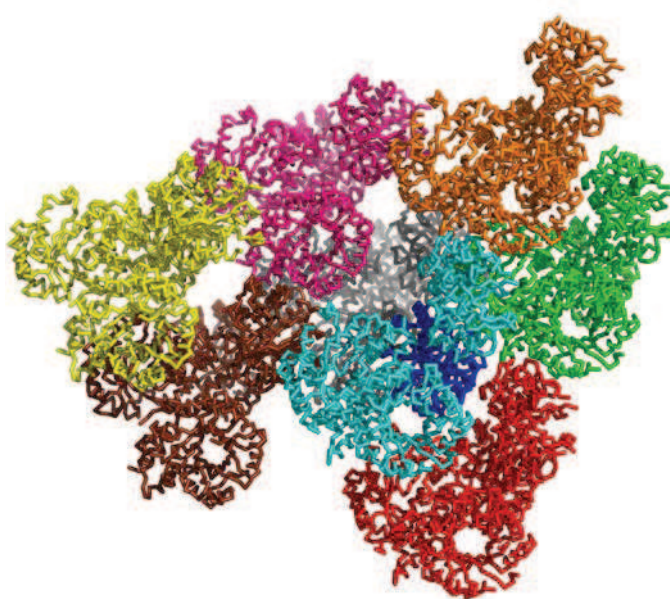


Figure I-18 Empilement cristallin de *MkHmdB*.

Une molécule d'HmdB est représentée en bleu foncé, ses 3 molécules voisines dans l'unité asymétrique en bleu ciel, et les autres groupes de 4 molécules en couleurs diverses.

Quelques boucles de résidus de la protéine HmdB ne sont pas visibles et se situent sur le haut du tonneau. Si l'on considère les molécules d'une unité asymétrique représentées en bleu dans la Figure I-18, la molécule en bleu foncé est voisine d'une molécule en gris d'une autre unité asymétrique (Figure I-19 (a), à gauche). Ces deux molécules se font face par le haut de leur tonneau. Nous voulions voir si les boucles d'HmdB non visibles pourraient s'intégrer dans cet empilement cristallin si elles étaient fixes. Les structures d'*EcBioB* (code PDB 1r30, (Berkovitch, Nicolet et al. 2004) et *TmHydE* (code PDB 3iiz, (Nicolet, Amara et al. 2009) ont été superposées sur les deux molécules d'HmdB considérées. Comme on peut le voir dans la Figure I-19 (a), à droite, les boucles de BioB ou HydE ont des mauvaises interactions entre elles.

La superposition des structures de BioB ou HydE sur celle des autres molécules d'HmdB de la même unité asymétrique (en bleu ciel), et sur les molécules leur faisant face, amène à l'observation de la même mauvaise interaction dans un cas. Dans les deux autres cas, l'orientation des molécules d'HmdB se faisant face n'est pas la même, et est représentée dans la Figure I-19 (b), à gauche : la mauvaise interaction entre les boucles n'est donc pas la même (Figure I-19 (a), à droite).

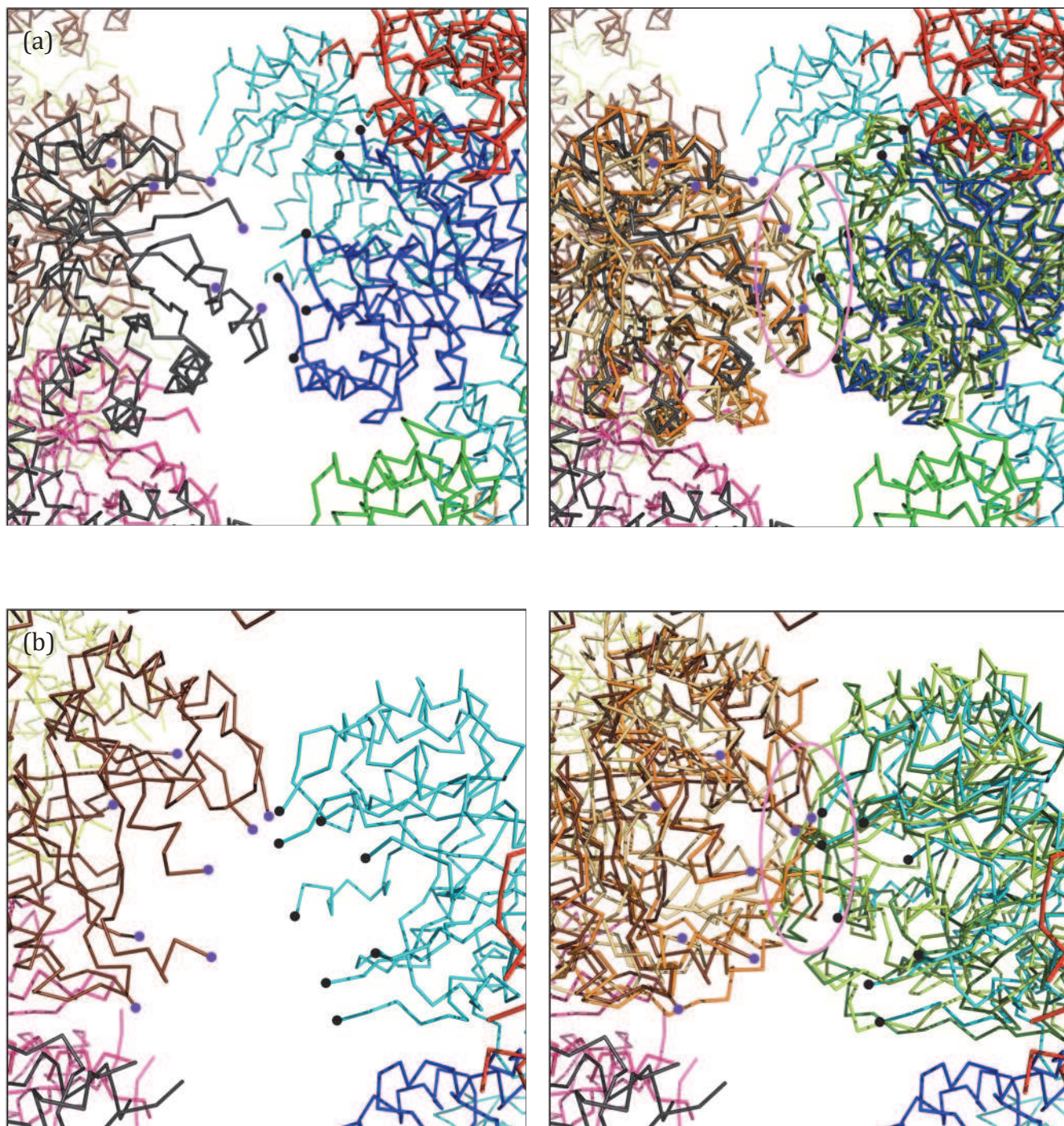


Figure I–19 Empilement cristallin de *MkHmdB*, boucles non visibles et superposition sur les structures de BioB et HydE.

Les résidus aux extrémités des boucles non visibles représentées par un cercle noir ou violet. A gauche, l'empilement cristallin de HmdB. A droite, les superpositions des structures de *EcBioB* (code PDB 1r30) et *TmHydE* (code PDB 3iiz) sur deux molécules différentes de *MkHmdB*. Deux formes de mauvaises interactions ont été observées et représentées en (a) et (b). Les couleurs utilisées pour représenter les molécules de *MkHmdB* sont les mêmes que dans la Figure I–18.

La superposition des structures de BioB ou HydE sur différentes molécules d'HmdB dans sa forme cristalline montre une non compatibilité de cet empilement cristallin avec la présence d'un centre [4Fe-4S] structurant les boucles du sommet du tonneau. Ceci est vrai pour chacune des quatre molécules de HmdB de l'unité asymétrique, et permet de comprendre pourquoi seul les conditions de cristallisation faisant appel à un agent chélateur de fer ont permis d'obtenir des cristaux. De plus, cette observation explique pourquoi les différents essais de reconstitution *in vitro* du centre [4Fe-4S] dans les cristaux ont échoué.

Les résidus présents à la surface d'HmdB sont en majorité de caractère polaire. La représentation de la surface de la protéine indique ce caractère (Figure I–20). L'utilisation d'additifs qui interagirait avec les molécules de protéine lors de la cristallisation a alors été testé mais n'a pas permis d'obtenir une nouvelle forme cristalline.

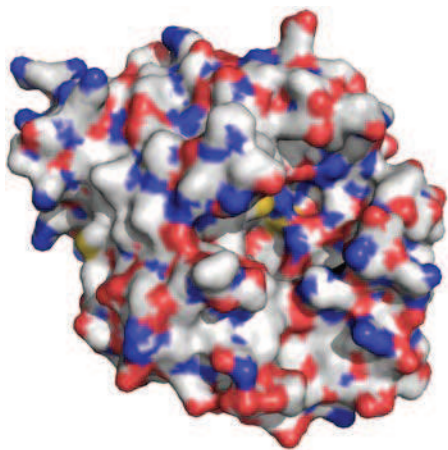


Figure I–20 Surface de *MkHmdB*.

Représentation de la surface de *MkHmdB* par coloration bleue des atomes d'azote, rouge des atomes d'oxygène, jaune des atomes de soufre et blanche des atomes de carbone.

I.6) Elaboration d'un mutant C-terminal et changement d'organisme

Malgré les informations intéressantes que pourrait nous fournir la structure cristallographique de HmdB, le fait que le centre [4Fe-4S], la SAM et les boucles qui constituent le sommet du tonneau sont absentes, ne permet pas une interprétation plus poussée. De plus, la recherche de ligands potentiels par accostage moléculaire, comme cela a été fait

pour la protéine HydE (Nicolet, Rubach et al. 2008), n'est pas vraiment possible. Il serait possible d'envisager de modéliser ces parties manquantes pour disposer d'un modèle complet raisonnable. Cependant, l'obtention d'une structure cristallographique serait préférable. Etant donnée la conclusion de l'analyse de l'empilement cristallin, une nouvelle forme cristalline est nécessaire. Par ailleurs, les essais récurrents de cristallisation n'ont pas permis d'obtenir celle-ci. Afin d'augmenter nos chances de succès, nous avons décidé de modifier l'objet à cristalliser. L'analyse de la structure de HmdB native montre que le groupe de résidus suivant l'hélice $\alpha 8$ (de la Gly336 à l'extrémité C-terminal) forme un ensemble flexible (car non visible) (Figure I-10), ce qui pourrait empêcher HmdB de cristalliser en augmentant l'hétérogénéité de l'échantillon. Ces résidus ne sont pas conservés (Figure I-17) entre les différents organismes, ils ne doivent donc pas avoir une fonction particulière. L'analyse des séquences de HmdB dans cette région nous a permis de retrouver un motif conservé GxxP qui suit l'hélice $\alpha 8$. Ce motif est aussi présent et conservé dans les protéines HydE et BioB. Il apparaît aussi dans les protéines HydG (Rubach, Brazzolotto et al. 2005), où il fait la jonction entre le domaine PRS et un domaine C-terminal responsable de la production de monoxyde de carbone (Nicolet, Martin et al. 2010). Il a d'ailleurs pu être montré que l'introduction d'un codon stop juste après ce motif GxxP permet d'obtenir une protéine HydG ne contenant que le domaine PRS, stable et active (Nicolet, Martin et al. 2010). Nous avons donc suivi une approche similaire en introduisant un codon stop à la place de la glycine 336 (Figure I-17). L'expression du mutant se fait tout aussi bien que la protéine native (Figure I-21). Le poids moléculaire de la protéine mutée est de 37.2 kDa (39.3 kDa pour la protéine native).

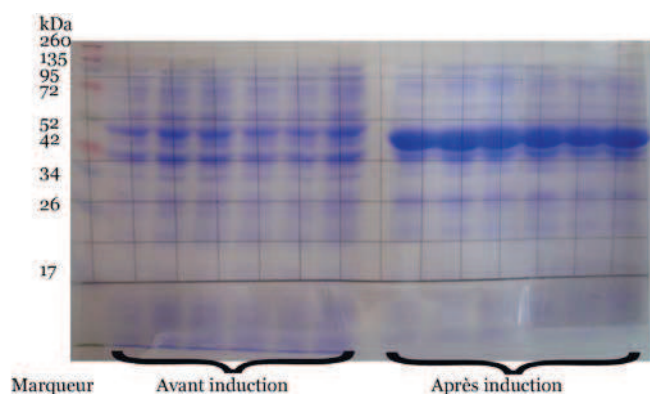


Figure I-21 Expression de Hmdb mutante.

SDS-PAGE d'expression de HmdB mutante. De gauche à droite : marqueur de poids moléculaire, 6 échantillons avant induction, 6 échantillons après induction. Notez l'apparition de la bande majoritaire à 40 kDa après induction.

La protéine se purifie de la même façon que la protéine native. Le profil du chromatogramme est le même (non montré) et la pureté de la protéine est moindre cette fois-ci par rapport à la protéine native (Figure I-22). Une étape de purification supplémentaire pourrait être ajoutée.

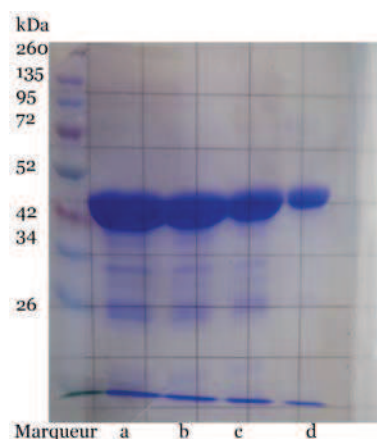


Figure I-22 Hmdb après purification.

SDS-PAGE de Hmdb après purification. De gauche à droite : marqueur avec les poids moléculaires des protéines standard indiquées, a : 10 μ g de protéine déposés, b : 7 μ g, c : 5 μ g, d : 2 μ g.

Le dosage en atomes de fer de l'échantillon indique la présence de 1.9 ± 0.2 atomes de fer par monomère de HmdB. Or nous attendons 4.0 atomes de fer par molécule de HmdB au minimum, comme dans le cas de la protéine native reconstituée (I.1)). L'analyse par spectroscopie UV-visible (Figure I-23) montre les mêmes pics que dans le cas de la protéine native : un pic à 280 nm, un épaulement à 320 nm et un pic large à 420 nm. Ces derniers sont caractéristiques de la présence de centres [4Fe-4S]. Comme dans le cas de la protéine native, malgré la surexpression de l'opéron ISC d'*E. coli* et la purification de la protéine en condition anaérobie, l'incorporation du centre [4Fe-4S] n'a pas été complète. Afin d'avoir un échantillon homogène de la protéine, la reconstitution *in vitro* du centre FeS a été faite (voir Matériel et Méthodes). L'échantillon après reconstitution contient alors 4.1 ± 0.4 atomes de fer par monomère de HmdB. L'analyse par spectroscopie UV-visible (Figure I-23) montre la présence des mêmes pics qu'avant reconstitution, caractéristiques de la présence de centres [4Fe-4S]. Le rapport DO à 280 nm / DO à 400 nm est plus faible, la reconstitution a donc augmenté la quantité de protéines contenant un centre [4Fe-4S].

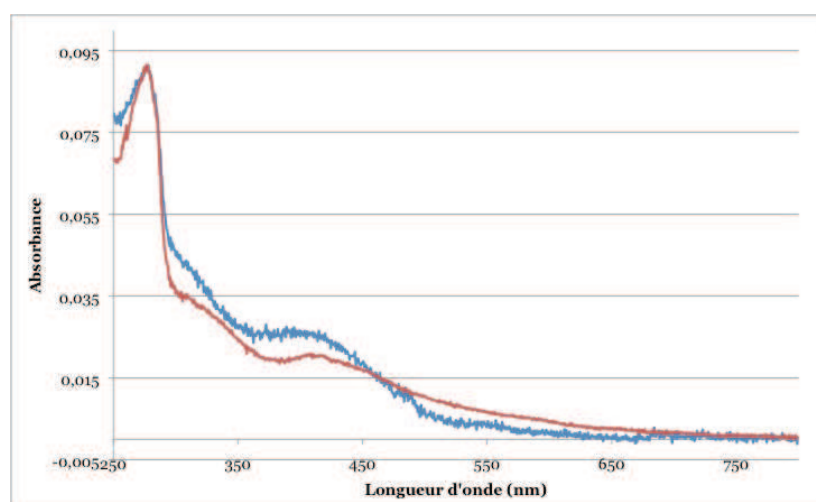


Figure I-23 Reconstitution de HmdB mutée.

Spectres UV-visible de la protéine avant (en rouge, courbe normalisée à 280 nm par rapport à la courbe bleue) et après reconstitution (en bleu).

Le spectre avant reconstitution a été fait à une concentration en protéine d'environ 0.13

mg/mL ; le tampon utilisé pour la dilution est celui dans lequel se trouve la protéine, c'est-à-dire le filtrat de la concentration de la protéine après la colonne Blue Sepharose (Tris 20mM pH 7.5, concentration non connue en NaCl). Le spectre après reconstitution a été fait à une concentration en protéine d'environ 0.09 mg/mL ; le tampon utilisé est celui dans lequel se trouve la protéine, c'est-à-dire le Tris 20mM pH 7.5, 300mM NaCl).

Des essais de cristallisation ont alors été faits en BAG avec le robot de cristallisation. 480 conditions ont été testées et n'ont abouti qu'à des cristaux dans les conditions contenant du citrate.

Le fait d'enlever la boucle de l'extrémité C-terminale ne modifie donc pas suffisamment l'objet pour permettre une cristallisation différente. Cette boucle n'est peut-être même pas responsable, à elle seule, de la difficulté de cristalliser une forme d'HmdB contenant son centre [4Fe-4S] et la SAM. Une autre façon de modifier l'objet à cristalliser consiste à utiliser une protéine orthologue issue d'un autre organisme. Ainsi, les acides aminés qui constituent la surface de la protéine seront différents et pourront engendrer des contacts cristallins différents. Nous avons choisi la protéine HmdB issue de *Methanococcus jannaschii* (MjHmdB). MjHmdB (locus JCVI : MJ_0785, code d'accès Uniprot Q58195) possède 375 acides aminés et un poids moléculaire de 42.8 kDa. Son pI de 8.65. Elle partage 55% d'identité de séquence avec son orthologue de *Methanopyrus kandleri*.

Le SDS-PAGE d'expression de MjHmdB (Figure I-24) montre l'apparition d'une bande dans les échantillons après induction correspondant à un poids moléculaire d'environ 45 kDa. Cette bande est moins intense qu'elle l'était pour l'expression des protéines MhHmdB native et mutante. Le taux d'expression de MjHmdB semble donc être moindre.

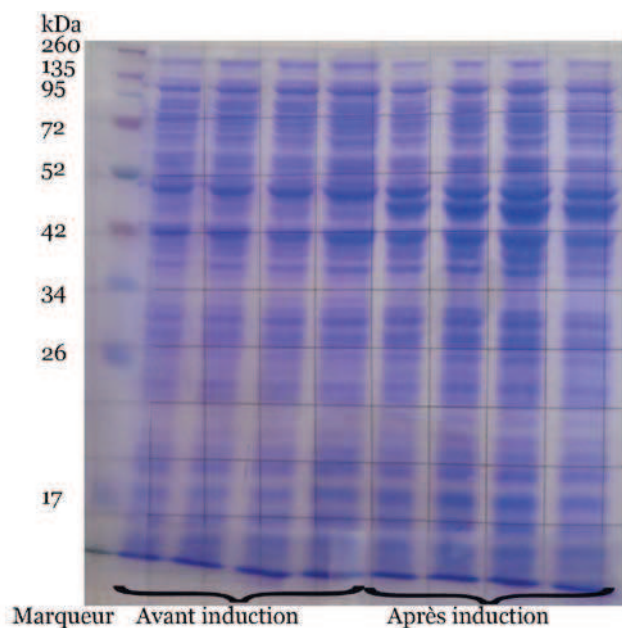


Figure I-24 Expression de HmdB de *Methanococcus jannaschii*.

SDS-PAGE d'expression de la protéine HmdB de *Methanococcus jannaschii*. De gauche à droite : marqueur protéique, 6 échantillons avant induction, 6 échantillons après induction.

Un premier protocole de purification de la protéine a été défini (voir Matériel et Méthodes). La première étape du protocole correspond à la lyse des cellules par ultrasons en présence de Benzonase nucléase. Pour la seconde étape nous avons choisi d'utiliser la propriété de thermostabilité de *MjHmdB* en chauffant l'échantillon à 70°C pendant 20 min. A cette température les protéines contaminantes d'*E. coli*, organisme de surexpression de la protéine d'intérêt, sont dénaturées et peuvent être éliminées par centrifugation. Ces deux premières étapes sont identiques à celles des protocoles de purification de *MkHmdB* native et mutante. Pour la troisième étape, une colonne échangeuse de cations, la résine SP Sepharose Fast Flow (purification à pH 6.5), a été choisie étant donné que le point isoélectrique de cette protéine est de 8.65 (contre 5.72 pour la protéine de *Methanopyrus kandleri*). Les protéines acides, qui sont majoritaires, ne se lieront pas à la colonne puisqu'elles seront chargées négativement à ce pH. Nous avons préféré cette colonne à la colonne Blue Sepharose car nous espérions obtenir une quantité plus importante de protéine pure.

Le premier essai d'utilisation de ce protocole a abouti à une non élution de la protéine de la colonne SP-Sepharose. L'étape de chauffage semble avoir été limitante : la protéine se serait dénaturée malgré sa stabilité théorique. Un second essai a été fait, sans chauffage de l'échantillon. La protéine a élué de la colonne SP-Sepharose et un SDS-PAGE a été fait sur les différents pics (non montré). L'analyse de ce gel montre que la protéine s'est retrouvée dans seulement quelques fractions, en faible quantité. Elle a malheureusement précipité au bout de quelques heures. Le pH de 6.5 semble déstabiliser la protéine. La purification à des valeurs de pH plus élevées pourrait donc être essayée. Cependant le pH ne peut pas être trop augmenté pour la colonne SP-Sepharose car le pI calculé de *MjHmdB* est de 8.65. Dans le cas de changement de valeur de pH, l'étape de chauffage pourrait être à nouveau intégrée dans le protocole, car la protéine sera peut-être plus stable.

Un autre protocole pourrait être défini en remplaçant la colonne SP-Sepharose par une colonne Blue Sepharose. Le pH utilisé pourrait être un peu plus élevé (par exemple 7.5).

I.7) Conclusion et perspectives

La protéine HmdB de *Methanopyrus kandleri* a été exprimée et purifiée. Elle contient 4 atomes de fer par molécule (même observation faite par (McGlynn, Boyd et al. 2010)). Son profil du spectre UV-visible est typique des centres [4Fe-4S].

Nous avons mené des essais de cristallisation afin de déterminer sa structure. Une forme de cristaux non colorés a été obtenue. Les conditions de cristallisation contenaient du tampon

citrate, chélateur de fer. Le remplacement de celui-ci par un tampon de même zone de virage n'a pas abouti à l'apparition de cristaux. Nous avons collecté un jeu de données sur cette forme de cristaux. La résolution de la structure a été faite par remplacement moléculaire et a nécessité d'élaborer un modèle de la protéine qui est une chimère des protéines BioB et HydE. Des cycles de modification de densité, correction du modèle et remplacement moléculaire ont été nécessaires avant de pouvoir affiner la structure avec le logiciel REFMAC5 (Murshudov, Vagin et al. 1997).

Les cristaux contiennent 4 monomères d'HmdB par unité asymétrique. Différentes boucles du haut du tonneau ne sont pas visibles, dont la boucle contenant les 3 cystéines conservées dédiées à la fixation du centre [4Fe-4S]. Le centre [4Fe-4S] n'est pas présent (cristaux non colorés), ainsi que la SAH (analogue de la SAM ajouté à la protéine).

La conformation de la protéine ne semble pas perturbée du fait de l'absence de la SAH, du centre [4Fe-4S] et de certaines boucles du haut du tonneau, mis à part pour les résidus 274 à 284. Les motifs conservés des protéines à radical SAM sont orientés de la même façon que dans les protéines BioB ou HydE.

Malheureusement cet empilement cristallin n'est pas compatible avec la présence des boucles fixes, d'après la superposition des molécules BioB ou HydE sur les molécules HmdB dans son empilement cristallin. Un mutant a été élaboré afin d'enlever les résidus de la Gly336 à l'extrémité C-terminal qui ne sont pas visibles pas dans la structure, donc flexibles. Ils pourraient adopter des conformations différentes à l'origine d'une non-homogénéité de l'échantillon. Les essais de cristallisation avec cette protéine ont malheureusement abouti à la même forme de cristaux.

Afin de modifier l'objet à cristalliser d'une manière plus conséquente, nous avons décidé de purifier la protéine HmdB de *Methanococcus jannaschii*. La protéine s'est révélée instable lors des premiers essais de purification. D'autres essais de purification sont à envisager.

Un modèle de la protéine apo-HmdB a été obtenu. Les parties manquantes suivantes de la protéine pourraient être modélisées en utilisant les structures de BioB et HydE :

- les résidus 64 à 87, le centre [4Fe-4S] et la SAH,
- les résidus 240 à 248, 303 à 315 (la région du résidu 336 au C-terminal n'étant pas importante), ainsi que la boucle 274 à 284.

Avec le modèle ainsi construit nous pourrions faire de l'accostage moléculaire de substrats d'HmdB comme il l'a été fait pour la protéine HydE (Nicolet, Rubach et al. 2008). L'accostage moléculaire est une méthode qui prédit l'orientation préférentielle d'une molécule

lors de sa fixation dans une protéine. Cette méthode permet aussi d'évaluer son affinité pour la protéine cible.

II) Etude structurale du complexe IscS-IscU

Malgré leur apparente simplicité, les centres FeS exclusivement composés d'atomes de fer et soufre sont synthétisés par des machineries protéiques.

Il existe trois systèmes bactériens d'assemblage des centres FeS : NIF, SUF et ISC (Johnson, Dean et al. 2005). Dans le cas d'ISC, la cystéine désulfurase (IscS), apporte le soufre nécessaire à la fabrication d'un centre [2Fe-2S] ou [4Fe-4S] (Agar, Krebs et al. 2000). IscU est la protéine échafaudage sur laquelle s'assemble le centre FeS. Elle possède trois cystéines et un acide aspartique conservés ainsi qu'une histidine semi-conservée dans son site d'assemblage (Shimomura, Wada et al. 2008). Les centres FeS assemblés *in vitro* dans IscU ne sont pas stables (Agar, Krebs et al. 2000). Néanmoins, il a été montré que la mutation de l'acide aspartique conservé en alanine permet de les stabiliser (Unciuleac, Chandramouli et al. 2007), (Shimomura, Kamikubo et al. 2007).

Ainsi, un complexe stable (IscS-IscU)₂ d'*Azotobacter vinelandii* contenant un centre [2Fe-2S] résistant à l'oxygène a été obtenu par mutation de l'Asp39 conservé de IscU en alanine (Raulfs, O'Carroll et al. 2008). Dans le but d'obtenir la structure tridimensionnelle de ce complexe, cette même mutation a été reproduite dans IscU d'*Archaeoglobus fulgidus* VC-16 (Af) sur l'Asp35 correspondant, dans le laboratoire de notre collaborateur, le professeur Dennis Dean à Virginia Tech aux Etats-Unis. Ceci a abouti à l'obtention d'un complexe (IscS-IscU)₂ contenant un centre [2Fe-2S] stable, résistant lui aussi à la dégradation par l'oxygène. L'organisme hyperthermophile réducteur de sulfate *Archaeoglobus fulgidus* VC-16 (Stetter, Lauerer et al. 1987) a été choisi parce que ses protéines sont particulièrement stables. Le choix d'étudier des protéines issues d'organismes hyperthermophiles a été discuté pour la protéine HmdB de *Methanopyrus kandleri* et *Methanococcus janaschii* (cf début de la partie I)).

Des structures de IscS ((Cupp-Vickery, Urbina et al. 2003), (Kaiser, Clausen et al. 2000) et IscU (Liu, Oganessian et al. 2005), (Shimomura, Wada et al. 2008)) ont déjà été déterminées. La structure du complexe natif IscS-IscU d'*E.coli* dépourvu de centres FeS, a été publiée récemment (Shi, Proteau et al. 2010).

II.1) Gènes d'*Archaeoglobus fulgidus* codant pour IscS et IscU

La machinerie ISC d'*Azotobacter vinelandii* et *Escherichia coli* est codée par un opéron contenant les gènes des protéines IscR, IscS, IscU, IscA, HscB, HscA, Fdx et IscX (Figure II-1). La séquence en acides aminés d'IscS d'*E. coli* (locus b2530) a été utilisée pour effectuer la recherche de séquences homologues (BLAST, (Altschul, Gish et al. 1990) chez l'organisme *Archaeoglobus fulgidus*.

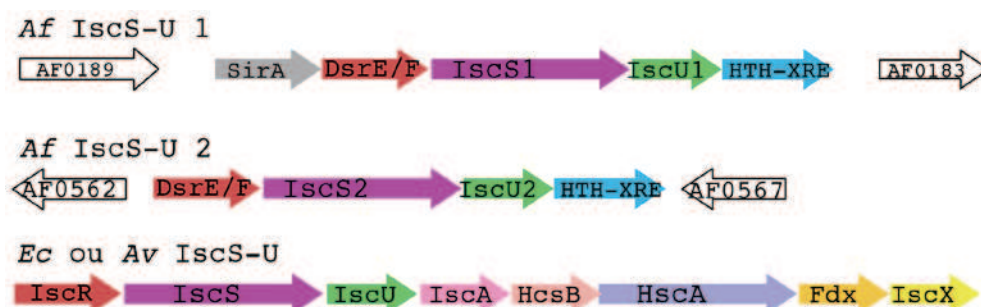


Figure II-1 Gènes d'*Archaeoglobus fulgidus* codant pour IscS et IscU.

De haut en bas :

Af IscS-U 1 : gènes IscS1 (AF0186) et IscU1 (AF0185) d'*Archaeoglobus fulgidus* et leurs voisins,
Af IscS-U 2 : gènes IscS2 (AF0564) et IscU2 (AF0565) d'*Archaeoglobus fulgidus* et leurs voisins,
Ec ou Av IscS-U : opéron ISC d'*E. coli* ou *Azotobacter vinelandii*.

Deux protéines possédant 40% d'identité de séquence avec IscS d'*E. coli* et 85% d'identité de séquence entre elles, ont été identifiées dans cette recherche. Elles sont chacune voisines (sites *Af IscS-U 1* et 2) de deux protéines qui possèdent 40% d'identité de séquence avec *EcIscU* (locus b2529) et 100 % d'homologie entre elles (Figure II-1). *A. fulgidus* ne contient pas d'homologues des protéines IscR, IscA, HscB, HscA, Fdx et IscX.

Les protéines codées par les gènes à proximité de ceux d'*AfIscS* et IscU ont été identifiées en utilisant un outil de recherche de domaines protéiques conservés (CDD, (Marchler-Bauer, Lu et al. 2011). Il s'agit de protéines homologues à DsrE/F (88% d'identité entre celles d'*AfIscS-U 1* et 2) et HTH-XRE (87% d'identité entre celles d'*AfIscS-U 1* et 2) (Figure II-1). DsrE/F est une protéine impliquée dans le métabolisme d'oxydation du soufre intracellulaire et fait partie d'un mécanisme relais de soufre pour la thiolation d'ARNt, tout comme son homologue TusD chez *E. coli* ((Dahl, Schulte et al. 2008), (Pott and Dahl 1998), (Numata, Fukai et al. 2006)). HTH-XRE (protéine XRE de structure hélice-coude-hélice (Luscombe, Austin et al. 2000)) est un régulateur de transcription (Wood, Devine et al. 1990). Sur le site *AfIscS-U 1* se trouve en plus un gène codant pour un homologue de la protéine SirA qui, entre autre, active des gènes de virulence (Goodier and Ahmer 2001).

II.2) Structure de *AflscS*

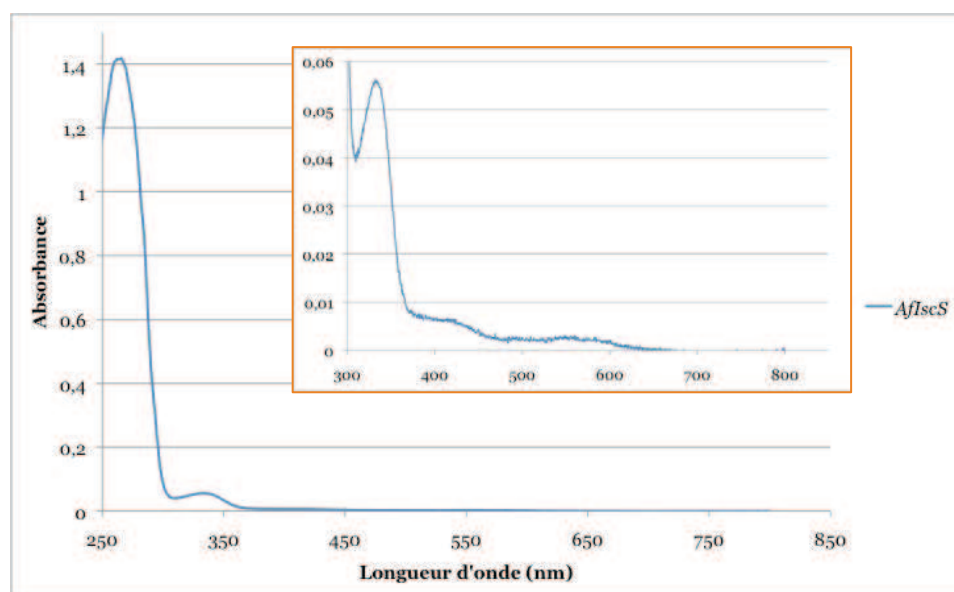
II.2.1) Caractérisation de *AflscS*

Les cystéines désulfurases catalysent la réaction (PLP)-dépendante de persulfuration d'une cystéine conservée par abstraction de l'atome de soufre d'une L-cystéine libre. Les produits de la réaction sont la L-alanine et l'enzyme persulfurée (Figure II–1 dans l'introduction).

Le site de fixation du PLP des cystéines désulfurases classiques comporte une lysine qui forme une base de Schiff avec ce dernier. Dans le cas d'*AflscS* la lysine est remplacée par un résidu acide aspartique (Figure II–6). A ma connaissance il n'y a pas d'autre cystéine désulfurase n'ayant pas cette lysine. La cystéine conservée se trouve sur une boucle labile, qui permet son mouvement depuis le site de fixation du PLP où a lieu la réaction de persulfuration vers des protéines clientes où l'atome de soufre est transféré.

Le gène AF0564 codant pour une cystéine désulfurase *AflscS* a été cloné dans le vecteur d'expression pAra13. La protéine a été produite et purifiée au laboratoire collaborateur du projet (voir Matériel et Méthodes). Le contenu en molécules de PLP par monomère d'IscS est de 0.27 (+/- 0.04) déterminée selon une méthode déjà décrite dans la littérature (Wada and Snell 1962).

Le spectre UV-visible d'une cystéine désulfurase classique est constitué de deux pics à environ 280 et 390 nm (Zheng, White et al. 1993), (Kaiser, Clausen et al. 2000), (Behshad, Parkin et al. 2004). Dans le cas d'*AflscS* (Figure II–2) il y a deux pics à 280 et 330 nm ainsi qu'une faible bande à 400 nm.



50mM pH 8, 400mM NaCl

Figure II–2
Spectre UV-visible de *AflscS*.
Spectre UV visible (zoom après 300nm dans l'encadré orange). Le spectre a été fait à une concentration en protéine d'environ 1.1 mg/mL ; le tampon utilisé est celui dans lequel se trouve la protéine, c'est-à-dire le Tris

La solution contenant *AfIscS* est incolore, alors qu'elle est jaune dans le cas d'une cystéine désulfurase classique. Cette différence est due au fait que la lysine qui forme une base de Schiff avec le PLP n'est pas présente dans *AfIscS*. L'absence d'une liaison aldimine entre la lysine et le PLP modifie l'absorbance UV-visible observée.

Mesurer l'activité cystéine désulfurase revient à mesurer la formation de produit au cours du temps, c'est-à-dire la L-alanine ou les ions sulfure. Des activités spécifiques de 67.6 nmoles de produit (L-alanine ou ions sulfures)/min/mg d'*AvIscS* (Zheng, Cash et al. 1998) et 240 nmoles de S^{2-} /min/mg d'*EcIscS* (Zhang, Urban et al. 2010) ont été mesurées. L'activité cystéine désulfurase d'*AfIscS* est de 0.00173 nmoles d'alanine/min/mg d'enzyme ce qui est donc très faible (tests d'activité effectués par le laboratoire collaborateur). L'ajout de PLP à la protéine n'augmente pas l'activité.

Il a été observé pour la cystéine désulfurase SufS (de la machinerie SUF) que sa faible activité était multipliée par 30 en présence de son partenaire protéique SufE (Ollagnier-De-Choudens, Lascoux et al. 2003).

Du fait de sa faible activité, il est possible qu'*AfIscS* possède aussi un partenaire protéique *in vivo* qui l'augmenterait. Le complexe *Af(IscS-Ala³⁵-IscU)₂* exprimé dans *E.coli* étant chargé en centres [2Fe-2S], ce partenaire aurait alors un homologue dans cet organisme.

Il est aussi possible que l'homologue de la protéine DsrE/F qui a un homologue TusD dans *E. coli* (voir paragraphe précédent) soit le donneur de soufre à *IscS* qui le donnerait alors à *IscU*.

II.2.2) Cristallisation de *AfIscS*

Les cristaux d'*AfIscS* (Figure II–3) ont été obtenus manuellement par la technique de la goutte suspendue. Les gouttes ont été préparées en ajoutant 3 μ L d'une solution d'*IscS* à 45 mg/mL à 1 μ L de solution réservoir (500 μ L) contenant 0.2 M acétate de calcium, 0.1 M cacodylate de sodium pH 6.5, 40% (v/v) PEG 300. L'expérience a été réalisée à 22-23° C chez nos collaborateurs et à 20° C dans notre laboratoire. Les cristaux apparaissent en 2 jours et sont congelés sous azote liquide sans cryoprotectant.

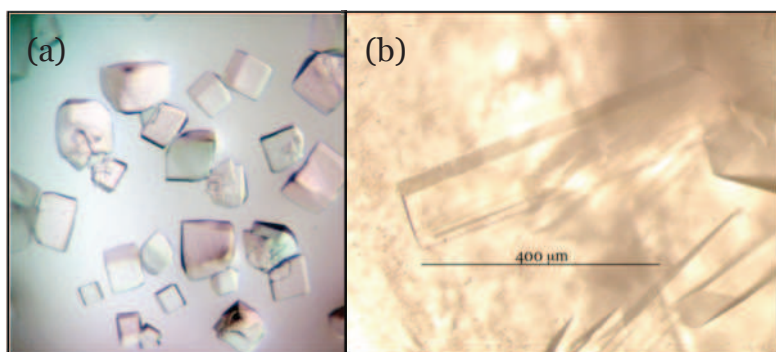


Figure II–3 Cristaux de *AfIscS*.

(a) Premiers cristaux de *IscS* obtenus par nos collaborateurs.
(b) Cristaux reproduits dans notre laboratoire. Les faciès des cristaux sont différents, ceci est probablement dû à des petites variations de conditions expérimentales.

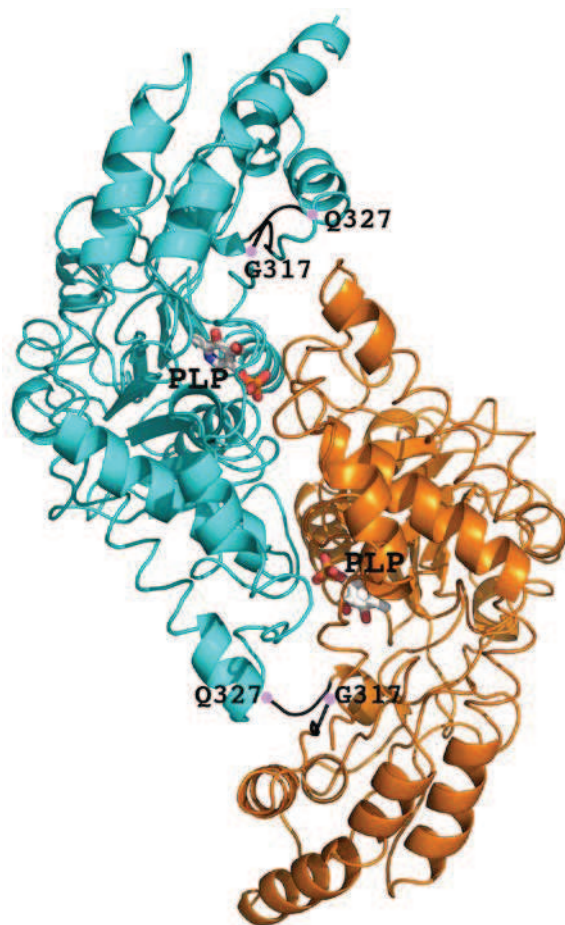
différences dans leurs cinétiques de croissance.

II.2.3) Structure de *AflscS*

Nous avons enregistré un jeu de données sur la ligne ID14eh4 à l'ESRF, Grenoble, à partir des premiers cristaux d'*AflscS* obtenus (Tableau II–1). L'intégration et la mise à l'échelle des données ont été faites à l'aide du programme XDS (Kabsch 2010). La structure de *AflscS* a été résolue à 1.43 Å de résolution par remplacement moléculaire avec le programme PHASER, en utilisant la protéine homologue d'*E.coli* (code PDB 1P3W) comme modèle de départ. Une seule solution a été trouvée par le programme avec une molécule d'IscS par unité asymétrique (résultat attendu par le calcul des coefficients de Matthews (Matthews 1968)). La haute résolution du jeu de données a permis de construire le modèle avec le logiciel ARP/wARP (Langer, Cohen et al. 2008), de l'améliorer avec le logiciel graphique COOT (Emsley, Lohkamp et al. 2010) et de l'affiner avec le logiciel REFMAC5 (Murshudov, Vagin et al. 1997).

Collecte et traitement de données				Affinement	
Ligne de lumière	ESRF ID14-eh4			R _{work}	0.112
Longueur d'onde (Å)	0.9395			R _{free}	0.152
Groupe d'espace	C222 ₁			Nombre d'atomes utilisés dans l'affinement (non hydrogène)	3570
Paramètres de maille (Å)	70.06	99.27	108.86	Molécules d'eau	381
Résolution (Å)	1.43			Ramachandran	
hkl observés	460751 (18683)			régions les plus favorables (%)	91.5
hkl uniques	68979 (4323)			régions permises supplémentaires (%)	8.5
Redondance	6.7 (4.3)			régions généreusement permises (%)	0
I/σ	24.08 (3.37)			régions interdites (%)	0
Complétude (%)	98.3 (84.4)			Déviati	
R _{sym} (%)	4.5 (42.0)			on “R.m.s.”	
				Liaisons (Å)	0.025
				Angles (°)	2.186

Tableau II–1 Statistiques de collecte, traitement de données et d'affinement pour *AflscS*.



AfIscS est un homodimère cristallographique (Figure II-4), comme dans le cas des structures d'autres cystéines désulfurases (Cupp-Vickery, Urbina et al. 2003), (Kaiser, Clausen et al. 2000), (Tirupati, Vey et al. 2004), (Shi, Proteau et al. 2010), ou en solution (Prischi, Pastore et al. 2010), (Agar, Zheng et al. 2000), (Zheng, White et al. 1993). Chaque monomère contient une molécule de PLP.

Figure II-4 Structure de *AfIscS*.

Les deux monomères d'*AfIscS* sont représentés par une couleur différente. Les résidus visibles de la boucle contenant la cystéine conservée (Cys321) sont de couleur noire. Les derniers résidus visibles de la boucle représentés par un disque violet et numérotés. Le PLP est représenté en bâtons.

II.2.3.1) Boucle contenant la cystéine conservée

La boucle contenant la cystéine conservée Cys321 d'*AfIscS* n'est pas visible (Figure II-4) comme dans les structures de *IscS* et *Nif-S* déjà publiées ((Cupp-Vickery, Urbina et al. 2003), (Kaiser, Clausen et al. 2000), (Shi, Proteau et al. 2010)).

Nous avons utilisé le logiciel Dali-server (Holm and Rosenstrom 2010) pour déterminer le degré de similarité structurale d'autres cystéines désulfurases avec *AfIscS*. Les protéines *IscS* d'*E. coli* ((Cupp-Vickery, Urbina et al. 2003), code PDB 1p3w) (utilisée lors du remplacement moléculaire), « *NifS-like* » de *Thermotoga maritima* (ou *IscS*) ((Kaiser, Clausen et al. 2000), code PDB 1eg5), *SufS* de *Synechocystis* sp. PCC 6803 ((Tirupati, Vey et al. 2004), code PDB 1t3i) et *CsdB* (ou *SufS*) d'*E. coli* ((Lima 2002), code PDB 1kmj) ont la plus forte similarité de séquence avec *AfIscS*. La superposition de ces structures montrent qu'elles sont globalement similaires à celle d'*AfIscS* (Figure II-5 (a)). Il a été discuté dans l'introduction (II.2.2.1)) que l'on peut distinguer deux groupes de cystéines désulfurases selon leurs similarités de séquence, et qu'une des régions les différenciant est celle où se trouve la cystéine active conservée (cercle sur la Figure II-5 (a)). La boucle peut être délimitée par les flèches noires

sur la Figure II–5 (b) et Figure II–6. Les structures des deux protéines SufS sont fortement similaires dans cette région, leur cystéine étant visible. Dans le cas des protéines IscS les résidus précédant et suivant la cystéine ne sont pas visibles. La boucle est plus longue dans les protéines IscS : ceci peut être expliqué par le fait qu'elle doit pouvoir atteindre plusieurs sites de protéines cibles qui ne s'associent pas de la même façon à IscS (par exemple, IscU, TusaA (Shi, Proteau et al. 2010) ou CyaY (Prischi, Konarev et al. 2010)).

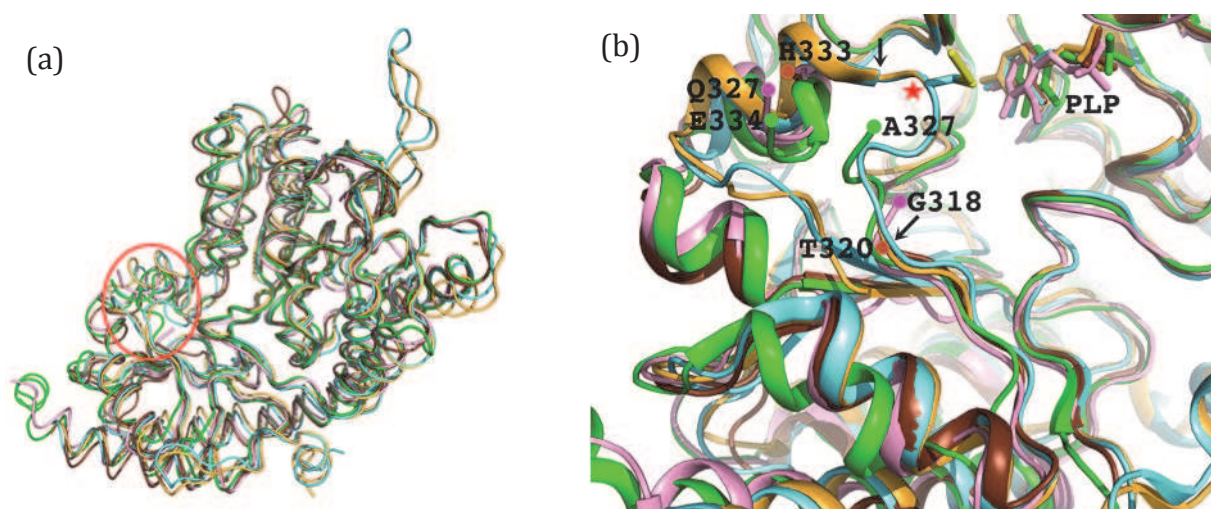


Figure II–5 Homologues d’AfIscS.

Superposition d’ IscS d’*E. coli* (code PDB 1p3w) (en vert), « NifS-like » (ou IscS) de *Thermotoga maritima* (code PDB 1eg5) (en brun), CsdB (ou SufS) d’*E. coli* (code PDB 1kmj) (en bleu), et SufS de *Synechocystis* sp. PCC 6803 (code PDB 1t3i) (en orange) sur AF0564 (*AfIscS*) (en rose).

(a) La région où se trouve la boucle dans le monomère de cystéine désulfurase est indiquée par un cercle rouge.

(b) Zoom sur la région des boucles qui contiennent la cystéine conservée (Cys321 pour AF0564). Les étoiles rouges indiquent les cystéines conservées visibles (Cys364 pour CsdB d’*Ec* qui est persulfurée dans cette structure, Cys372 pour SufS de *Synechocystis*). Dans les autres cas, les derniers résidus visibles de chaque extrémité de la boucle sont numérotés et sont indiqués par des cercles. Les flèches indiquent les résidus dont la position est indiquée dans la Figure II–6 afin de pouvoir comparer les boucles.

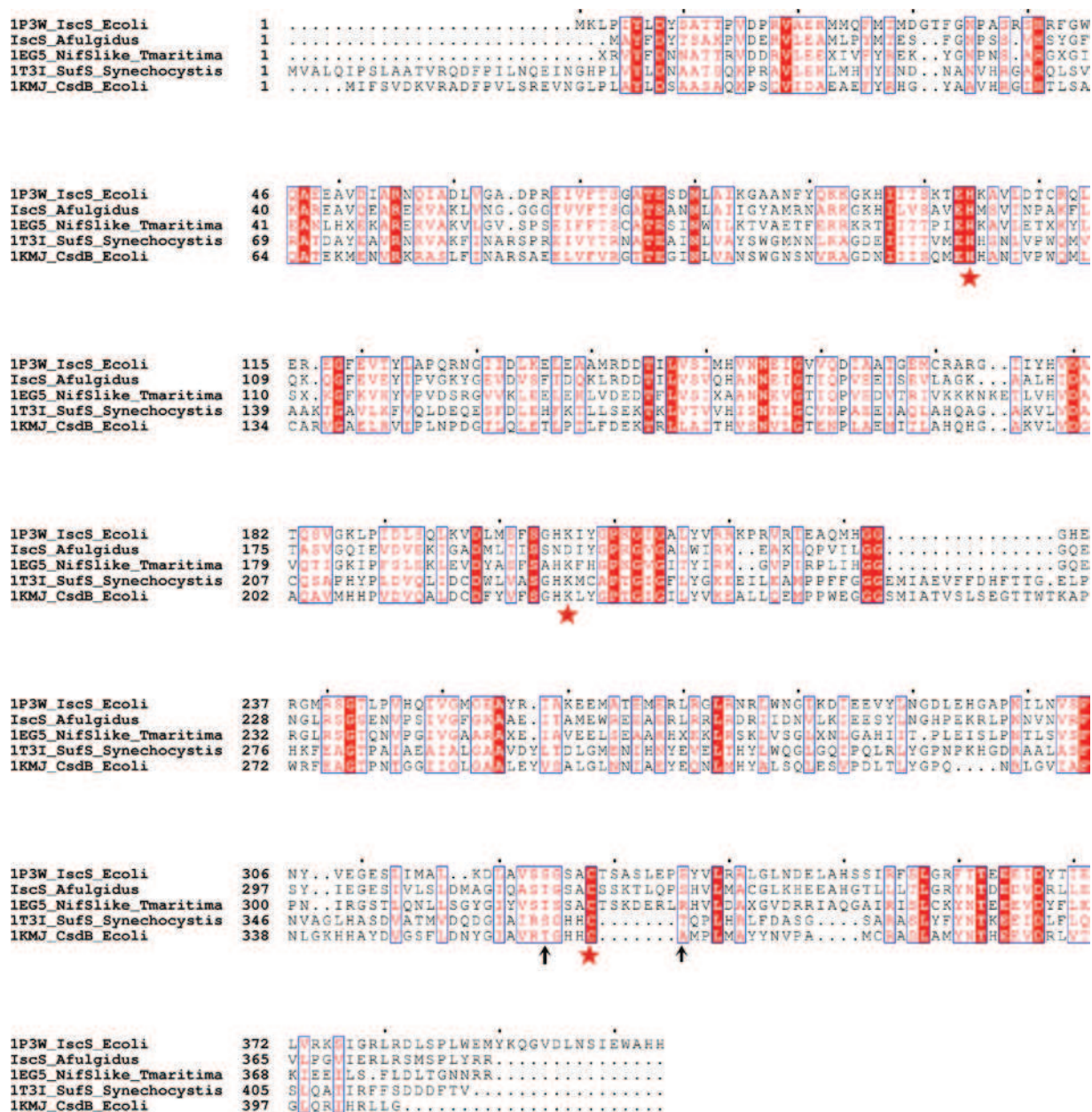


Figure II-6 Alignement de séquence multiple avec Clustal W2 (Larkin, Blackshields et al. 2007) de plusieurs IscS et homologues structuraux (Dali-server).

Alignement multiple de séquences d'IscS d'*E. coli* (code PDB 1p3w), AF0564 (*AfIscS*), « NifS-like » (ou IscS) de *Thermotoga maritima* (code PDB 1eg5), CsdB (ou SufS) d'*E. coli* (code PDB 1kmj), et SufS de *Synechocystis* sp. PCC 6803 (code PDB 1t3i) avec CLUSTALW2 (Larkin, Blackshields et al. 2007), corrigé manuellement au niveau de la région de la cystéine conservée pour tenir compte de la localisation structurale des résidus et de leur conservation spatiale. L'alignement de séquence est représenté avec le logiciel ESPript (Gouet, Courcelle et al. 1999) en utilisant la matrice de Risler et les paramètres par défaut. Les boîtes rouges et caractères blancs indiquent l'identité stricte, les caractères rouges indiquent la similarité dans un groupe et les cadres bleus indiquent la similarité à travers les groupes. Les étoiles rouges indiquent la position de la lysine semi-conservée formant une base de Schiff avec le PLP (Asp199 dans AF0564), de l'histidine conservée proche du PLP (His98 dans AF0564) et la cystéine conservée (Cys321 dans AF0564). Les flèches indiquent les résidus dont la position est indiquée dans la Figure II-5.

II.2.3.2) Site de fixation du PLP

Le PLP est présent (Figure II–7) dans la structure d'*AflscS* alors qu'il n'avait été ajouté ni lors de la purification, ni lors de la cristallisation. Il a une occupation de 0.7 dans les cristaux et de 0.27 en solution (cf II.2.1)). Il y a donc sélection par cristallisation des molécules d'*IscS* contenant du PLP. Dans son site de fixation, il y a interaction entre la pyridine de l'histidine conservée (His98) et le cycle du PLP. L'acide aspartique (Asp199) remplace la lysine qui forme une base de Schiff avec le PLP.

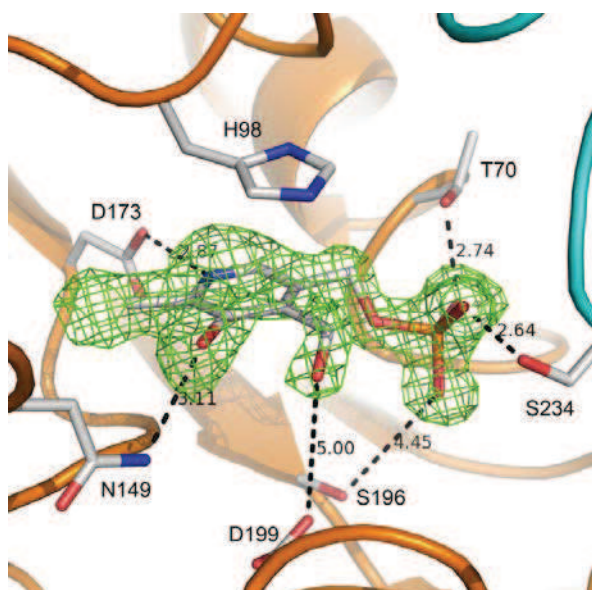


Figure II–7 Site de fixation du PLP dans *AflscS*.

Carte « omit » du PLP contournée à 2.5σ . Distances en Å.

Différents mécanismes réactionnels ont été proposés pour la cystéine désulfurase. Ils comportent quelques différences mais le principe reste le même (Zheng, White et al. 1994), (Kaiser, Clausen et al. 2000), (Behshad and Bollinger 2009). La L-cystéine se lie au site actif de l'enzyme. L'amine déprotonée de la L-cystéine fait une attaque nucléophile sur l'atome C4' du PLP pour produire une aldimine externe et le relargage de la chaîne latérale neutre de la lysine. Le proton en position α du substrat est éliminé pour produire un intermédiaire quinonoïde (quinonoid) (Figure II–8). Il a été proposé que soit la lysine, soit la cystéine du site actif ajoute un proton au C4' de l'intermédiaire quinonoïde, produisant l'intermédiaire kétimine (ketimine). La cystéine conservée déprotonée fait une attaque nucléophile sur l'intermédiaire kétimine, pour produire l'enzyme persulfurée et l'alanine α -énamine (Ala enamine). La protonation au C β du motif aminoacrylate résulte en la formation de l'intermédiaire alanine-ketimine, suivi par l'attaque nucléophile de la lysine déprotonée sur l'atome C4' du PLP, pour produire une aldimine interne.

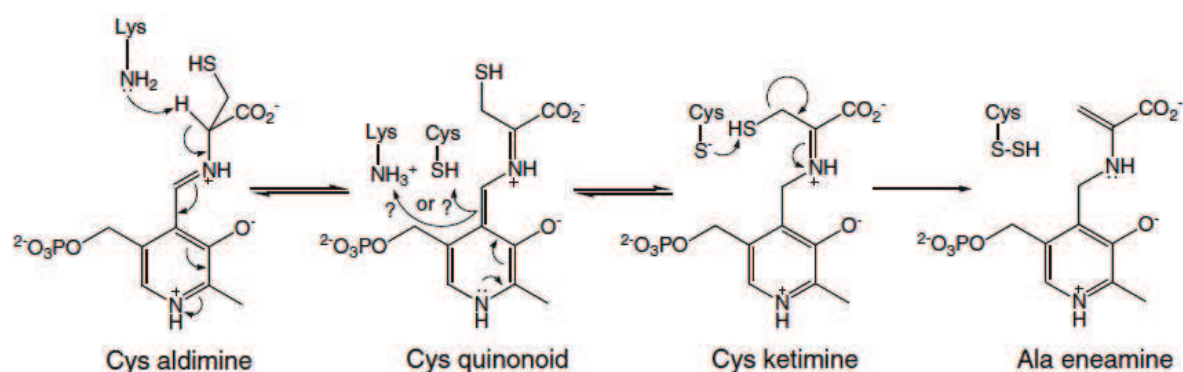


Figure II-8 Mécanisme réactionnel des cystéines désulfurase.

Figure tirée de (Hidese, Mihara et al. 2011).

Ces mécanismes font intervenir la lysine semi-conservée. Il serait donc intéressant d'étudier le mécanisme catalytique de cystéines désulfurases ne possédant pas cette lysine, telle qu'*AfIscS*.

Nous avons fait des essais de cristallisation de la protéine *AfIscS* en présence de L-cystéine, (en présence ou non de DTT afin d'empêcher la dimérisation de la cystéine en cystine) dans le but d'avoir la structure de la protéine avec son substrat. Ces essais n'ont pas abouti car nous n'avons pas vu la L-cystéine dans les structures obtenues.

II.3) Structure du complexe *Af(IscSU)*₂

II.3.1) Caractérisation de *AfIscSU*

Les gènes AF0564 et AF0565 codant pour *AfIscS* et *IscU* ont été clonés dans le vecteur d'expression pAra13. Le complexe protéique a été produit et purifié en condition anaérobie dans le laboratoire collaborateur du projet (voir Matériel et Méthodes II.2)).

Les dosages en fer et soufre du complexe indiquent la présence de 0.57 Fe (+/- 0.01) et 0.7 S (+/- 0.03) par molécule d'*IscU* (méthodes décrites dans (Raulfs, O'Carroll et al. 2008) et (Chen and Mortenson 1977)). Il y a donc plusieurs espèces en solution, avec et sans centres FeS. Le taux en PLP est le même que pour *IscS* seule (cf II.2.1)).

L'élution de la solution contenant le complexe *AfIscSU* sur une colonne d'exclusion stérique analytique Superdex 200 HR 10/30 produit deux pics (Figure II-9 (a)). D'après la courbe de calibration de la colonne, le premier pic, minoritaire, correspond à une masse moléculaire de 359 kDa et le deuxième pic à 133 kDa. Cette dernière équivaut à un hétérotétramère $\alpha_2\beta_2$, aussi observé chez *Azotobacter vinelandii* (Raulfs, O'Carroll et al. 2008). La valeur de 359 kDa du premier pic d'élution correspond approximativement à trois

hétérotétramères $\alpha_2\beta_2$. Un SDS-PAGE (non montré) confirme la présence exclusive d'espèces IscU et IscS dans les deux pics (bandes à 16.9 kDa et 41.8 kDa respectivement). Le premier pic représente donc un état d'agrégation de l'hétérotétramère et contient un léger épaulement qui correspondrait à une agrégation encore plus importante.

Nous avons alors étudié la nature de l'agrégation. Nous avons mis le complexe *AfIscSU* en présence d'un réducteur (1mM dithionite ou 5mM DTT) pendant 30 min. La solution a alors été injectée sur la même colonne d'exclusion stérique. Le profil d'élution n'est pas le même que dans le cas de la protéine native (Figure II-9). Lors du traitement au dithionite la taille du pic d'agrégation diminue d'environ 30%, tandis que lors du traitement au DTT il a disparu. Ces observations se retrouvent par analyse par gel natif (Figure II-10) : *AfIscSU* est caractérisée par une bande majeure qui correspond à l'hétérotétramère, une bande de poids supérieur (pic d'agrégation) et une bande minoritaire qui correspond à l'épaulement que l'on peut voir juste avant le pic d'agrégation. Lors du traitement au dithionite l'intensité de la bande correspondant au pic d'agrégation a diminué de 90% et celle de l'épaulement a disparu. Lors du traitement au DTT la bande du pic d'agrégation a complètement disparu.

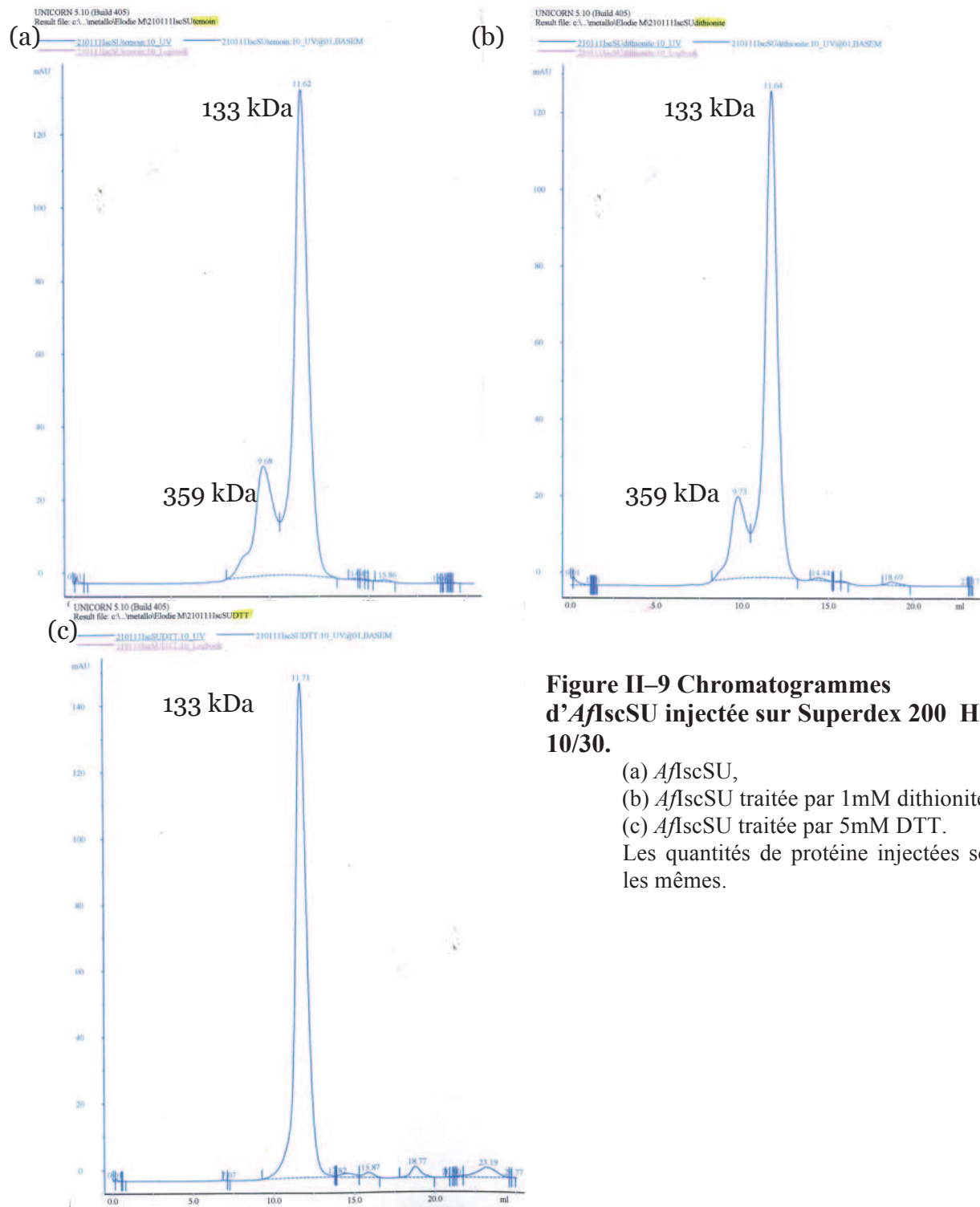


Figure II-9 Chromatogrammes d'*AflscSU* injectée sur Superdex 200 HR 10/30.

(a) *AflscSU*,
(b) *AflscSU* traitée par 1mM dithionite,
(c) *AflscSU* traitée par 5mM DTT.
Les quantités de protéine injectées sont les mêmes.

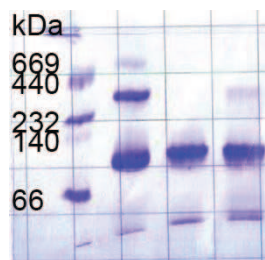
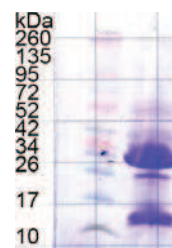


Figure II-10 Gel natif du complexe *AflscSU*.
De gauche à droite : marqueur (poids moléculaires indiquées sur la gauche du gel), *AflscSU*, *AflscSU* traitée au DTT et *AflscSU* traitée au dithionite. 20 μ g de protéine ont été déposés.

L'observation de l'effet de réducteurs sur l'agrégation de l'hétérotétramère a conduit à l'hypothèse de la présence de ponts disulfures intermoléculaires dans l'échantillon. Cette hypothèse a été étudiée par SDS-PAGE (non réducteur) (Figure II-11). Dans ces conditions non réductrices, le SDS devrait séparer les molécules IscS-IscU qui ne sont pas liées par des ponts disulfures, ce qui a été observé dans notre cas.

Figure II-11 SDS-PAGE (non réducteur) du complexe *AfIscSU*.

A gauche : marqueur (masses moléculaires indiquées sur la gauche du gel), à droite : *AfIscSU* (20 µg déposés).



L'agrégation d'IscSU ne résulterait donc pas de ponts disulfures. Le DTT est surtout employé en tant qu'agent réducteur des ponts disulfures intra ou intermoléculaires pouvant se former entre les résidus des cystéines de protéines. Un autre effet du DTT a été décrit par (Alliegro 2000) (Figure II-12) ; dans leur étude ces auteurs ont montré qu'il perturbe la fixation d'un ligand carbohydrate. En excès le DTT pourrait couvrir un domaine protéique en se liant à ses résidus glutamate ou aspartate ce qui pourrait empêcher la liaison du ligand.

Dans notre cas le DTT permet l'homogénéisation de l'échantillon.

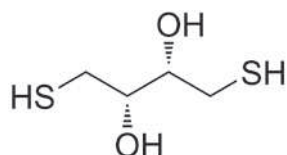


Figure II-12 Formule chimique développée du dithiothreitol.

Des spectres UV-visible ont été faits sur les échantillons avant et après traitement au DTT ou dithionite. Le spectre UV-visible du complexe *AfIscSU* avant traitement (courbes rouges sur la Figure II-13) montre les bandes caractéristiques de la présence de PLP à 330 nm (déjà observé pour IscS seule) et de centres [2Fe-2S] à environ 400 nm. Après traitement par 5 (ou 50 mM) de DTT ou 1 mM dithionite, le profil du spectre UV-visible ne change pas. De même, l'exposition à l'air après traitement ou non au DTT ne change rien au profil du spectre UV-visible, même après deux jours (Figure II-13). Le centre [2Fe-2S] est donc stable à l'air, comme observé pour *A. vinelandii* (Raulfs, O'Carroll et al. 2008).

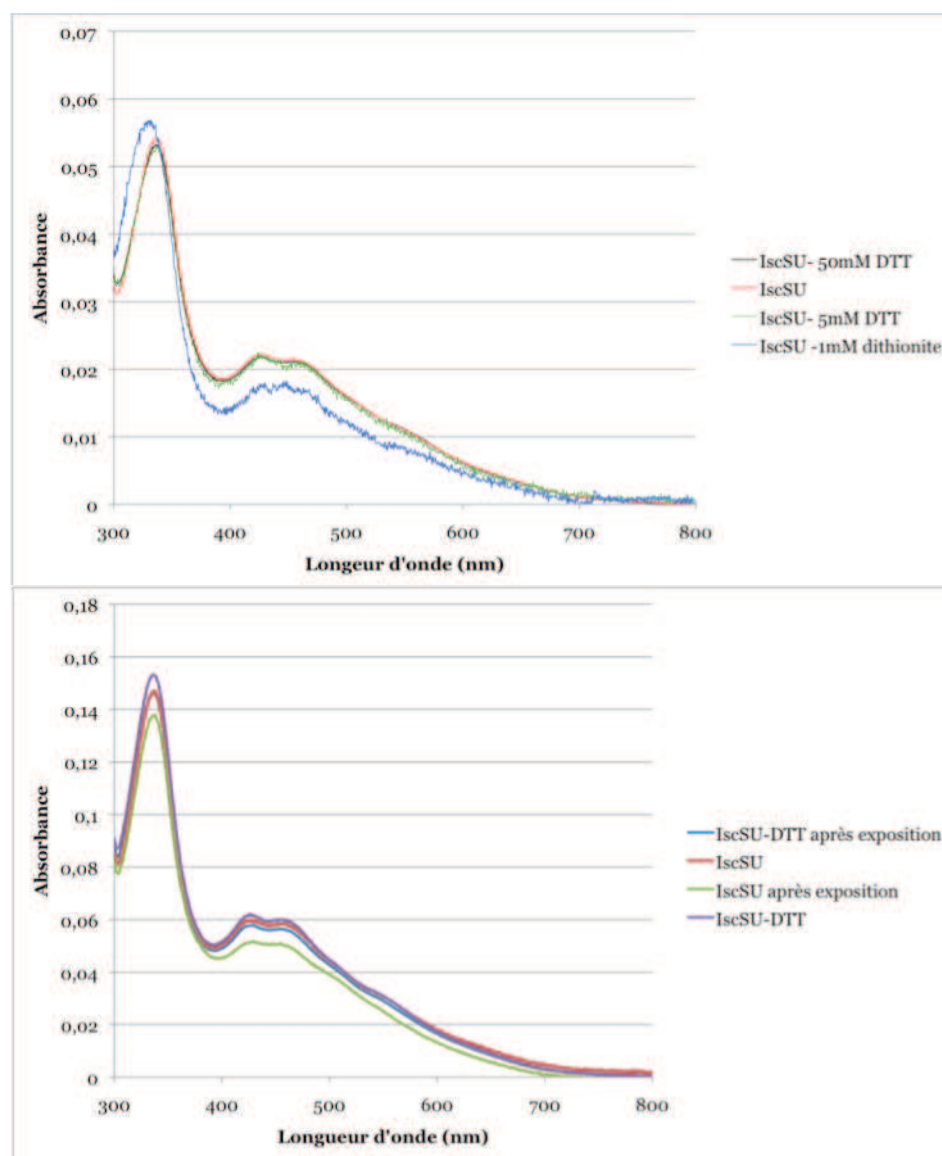


Figure II-13
Spectres UV-
visible d'AflscSU.

Spectres UV-visible du complexe AflscSU : en haut : AflscSU traitée par 5 ou 50 mM de DTT, par 1mM dithionite. NB : Les spectres de AflscSU non traitée, ou traitée par 5 ou 50 mM de DTT se superposent.

En bas : AflscSU traitée ou non par 5 mM de DTT puis exposée à l'air. La concentration en protéine des échantillons non traités au DTT était moins élevée que celle des autres échantillons.

Les courbes ont été normalisées par rapport à l'absorbance à 280 nm. La concentration en protéine dans les échantillons était

d'environ 0.75 mg/mL. Le tampon utilisé pour diluer la protéine est le 13mM Tris pH8 38 mM NaCl.

II.3.2) Cristallisation du complexe AflscSU

Après l'obtention du complexe AflscSU purifié, des essais de cristallisation ont été menés chez nos collaborateurs afin d'ensuite obtenir sa structure. La recherche des conditions de cristallisation a été faite sur la paillasse à l'aide d'un robot rapidement après avoir sorti la protéine de la boîte à gants (BAG). Plusieurs conditions ont permis d'obtenir des cristaux qui étaient tous de longues baguettes fines. Des essais de reproduction de ces cristaux ont été faits manuellement. La condition 0.2 M acétate d'ammonium, 0.1 M Hepes pH 7.5, 25% w/v PEG 3350 a alors été retenue car elle a permis de réobtenir des cristaux. La goutte de cristallisation a été préparée en ajoutant 2 μ L de protéine à 20.5 mg/mL (dans le tampon 13 mM Tris pH 8, 38 mM NaCl) à 2 μ L de solution réservoir. Le puits contenait 190 μ L de solution réservoir et

10 μ L d'eau afin de modifier la condition d'équilibration, qui fut le seul moyen pour nos collaborateurs d'obtenir des cristaux manuellement par la technique de la goutte assise. Des cristaux de couleur rouge-marron et 400 μ m de long sont apparus après 2 jours, et ont été congelés rapidement en utilisant la solution cryoprotectrice suivante : 0.2 M acétate d'ammonium, 0.1 M Hepes pH 7.5, 25% w/v PEG 3350, 25% v/v glycérol. Cette forme cristalline obtenue sur la paillasse sera appelée « exposée à l'air » (AE, en anglais).

Une autre forme cristalline a été obtenue dans notre laboratoire en BAG (à 20°C), en utilisant la même condition de cristallisation que la forme AE, avec la protéine préincubée dans une solution 5mM de DTT. Ce dernier a été utilisé après avoir observé son effet d'élimination de l'agrégation de la protéine. Le puits contenait 500 μ L de solution réservoir sans ajout d'eau. Ils seront appelés forme « traitée au DTT » (DT, en anglais). Les cristaux ont été obtenus en 2 jours et ont été congelés 3 jours après leur apparition. Les premiers cristaux ont été obtenus avec la technique de la goutte assise. L'essai de reproduction avec la technique de la goutte suspendue en utilisant des boîtes EasyXtal (132023, Qiagen, puits à visser) n'a pas abouti. En revanche, ils ont été obtenus par la suite dans des boîtes VDX prégraissées (HR3-170, Hampton Research, lamelles graissées pour permettre leur adhésion au puits). Il y a donc dans ce cas une importance de choix de matériau pour le support de gouttes, le premier étant en plastique et le second en verre.

Une troisième forme cristalline a été obtenue dans notre laboratoire en BAG dans les mêmes conditions que les cristaux de la forme DT mais avec la protéine purifiée en condition anaérobie sans traitement avec un réducteur. Ils seront appelés forme « non traitée » (UT, en anglais). Dans ce cas, le choix du matériau des boîtes de cristallisation s'est aussi révélé important car les cristaux n'ont pas été obtenus dans les boîtes EasyXtal.

Dans les trois cas les cristaux sont colorés, indiquant la présence de centres FeS (Figure II–14).

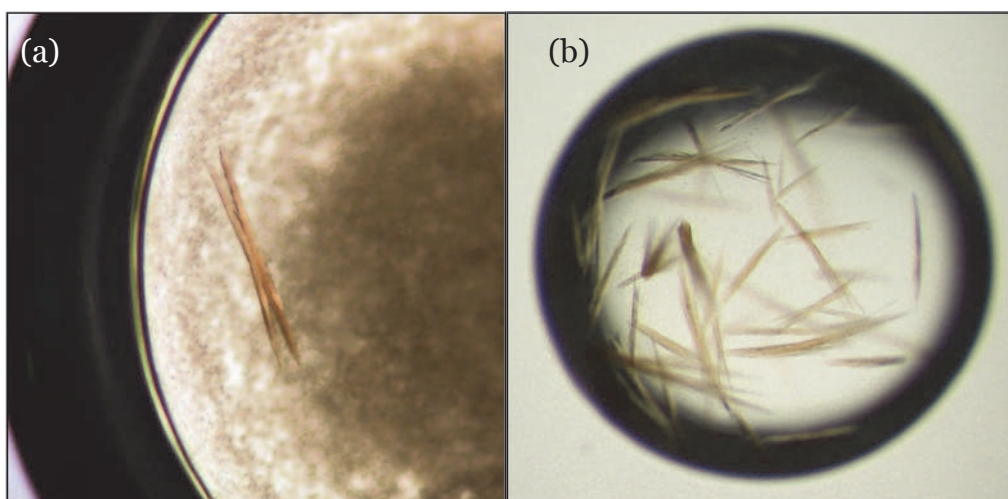


Figure II-14 Cristaux du complexe *AfIscSU*.

(a) Cristaux de IscSU exposé à l'air.

(b) Cristaux de IscSU traité au DTT (même type de cristaux pour la forme UT).

*II.3.3) Résolution de structure de *AfIscSU**

Des jeux de données ont été collectés pour les trois formes cristallines sur la ligne ID23-eh2 de l'ESRF. Les données ont été traitées avec le programme XDS (Kabsch 2010). Seules les formes AE et DT ont été affinées (Tableau II-2). La forme UT a seulement servi à confirmer que sa structure coïncide avec celle de l'état DT.

Collecte et traitement de données		
Protéine	AE IscUS	DT IscUS
Ligne de lumière	ESRF ID23-eh2	ESRF ID23-eh 1
Longueur d'onde (Å)	0.8726	0.9801
Groupe d'espace	$P2_12_12_1$	$P2_12_12_1$
Paramètres de maille (Å)	75.47 96.13 152.79	76.41 94.90 150.59
Résolution (Å)	2.75	2.53
hkl observés	154457 (21174)	137031 (18680)
hkl uniques	29148 (4486)	35945 (5446)
Redondance	5.3 (4.7)	3.8 (3.4)
I/ σ	14.76 (3.12)	13.18 (3.19)
Complétude (%)	97.5 (93.9)	96.2 (91.8)
R _{sym} (%)	7.7 (39.9)	7.8 (40.2)

Affinement		
Protéine	AE IscUS	DT IscUS
R_{work}	0.214	0.211
R_{free}	0.273	0.287
Nombre d'atomes utilisés dans l'affinement (non hydrogène)	6517	7633
Molécules d'eau	48	143
Ramachandran		
régions les plus favorables (%)	87.2	89.0
régions permises supplémentaires (%)	11.9	9.9
régions généreusement permises (%)	0.6	0.9
régions interdites (%)	0.3	0.2
Déviations "R.m.s."		
Liaisons (Å)	0.015	0.014
Angles (°)	1.641	1.623

Tableau II–2 Statistiques de collection, traitement de données et d'affinement pour le complexe *AfIscSU*.

La structure du complexe AE a été résolue à 2.75 Å de résolution par remplacement moléculaire avec le programme PHASER (McCoy, Grosse-Kunstleve et al. 2007), en utilisant la structure à haute résolution de *AfIscS* préalablement résolue au laboratoire et la structure de IscU d'*Aquifex aeolicus* (code PDB 2z7e) (Shimomura, Wada et al. 2008) comme modèles de recherche (2 copies de chaque, selon le calcul des coefficients de Matthews (Matthews 1968)). Une seule solution a été trouvée par le programme avec deux molécules d'*AfIscS* et d'*AfIscU* par unité asymétrique. La construction a été faite avec le logiciel graphique COOT (Emsley, Lohkamp et al. 2010) et l'affinement avec les logiciels AUTOBUSTER (Blanc, Roversi et al. 2004) et REFMAC5 (Murshudov, Vagin et al. 1997). Une des molécules de IscU (molécule D) étant partiellement désordonnée, elle n'a pas été incluse dans le modèle.

II.3.4) Structure de IscSU exposée à l'air (AE)

Le complexe *AfIscSU* est un hétérotétramère $\alpha_2\beta_2$, avec un dimère d'IscS et deux monomères d'IscU qui se lient près des extrémités C-terminal des molécules d'IscS, formant complexe en forme de S (Figure II–15). Cette forme est similaire à celle du complexe apo d'*E. coli* (Shi, Proteau et al. 2010).



Figure II–15 Hétérotétramère du complexe AE-*Af(IscSU)*₂.

En rose: IscU. En orange et vert : IscS. La position approximative de la molécule D partiellement désordonnée est indiquée par un cercle violet. Les molécules de PLP sont représentées par des bâtons et sont indiquées par des flèches noires. Les cystéines conservées des protéines IscS sont représentées par des bâtons et sont indiquées par des flèches violettes.

La boucle contenant la cystéine conservée d'IscS (Cys321) est visible (molécule A) (Figure II–16). La Cys321 fait face aux 3 cystéines conservées Cys33, Cys58 et C102 d'IscU (molécule C), une histidine semi-conservée (His101) et l'Ala35. Cette dernière résulte de la mutation Asp→Ala discutée plus haut. La boucle contenant la Cys321 est partiellement désordonnée, la Lys324 n'est pas visible. La boucle contenant la Cys33 est aussi partiellement désordonnée, son atome de soufre n'est pas visible et n'a donc pas été inclus dans le modèle. Les 3 cystéines conservées de IscU et la cystéine conservée d'IscS font face à un pic positif de densité électronique d'intensité 7σ dans la carte du type (Fo-Fc) (Figure II–16 A) indiquant la présence d'un centre FeS. Celui-ci ne semble pas être un agrégat [2Fe-2S] classique, vu la forme du pic. Lorsqu'on calcule une carte de densité électronique avec des coefficients (Fo-Fc) après avoir enlevé du modèle les atomes (carte « omit ») du centre FeS ainsi que le C β , et le S γ de la Cys321, les deux régions Cys321/centre FeS semblent être reliées (Figure II–16 B).

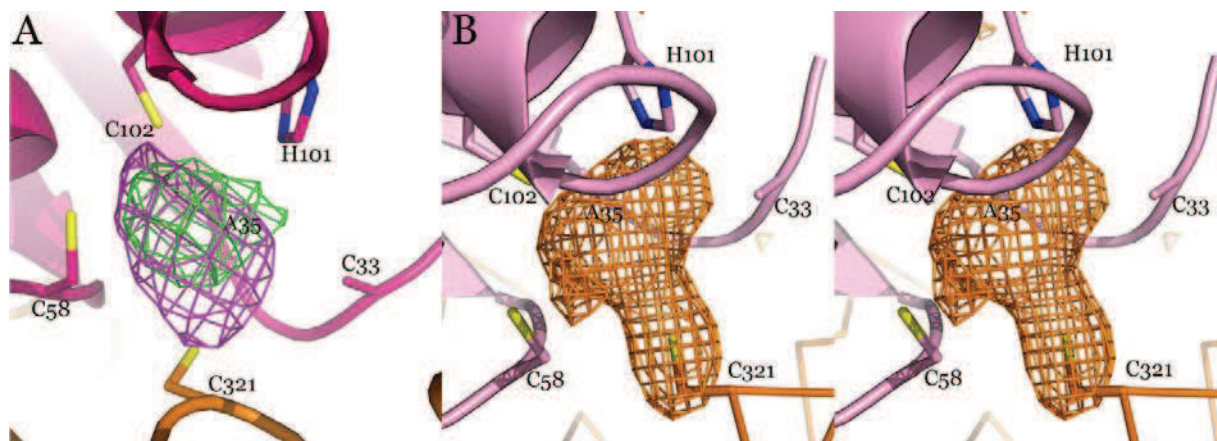


Figure II–16 Site d’assemblage du centre FeS dans IscU dans le complexe AE *AflscSU*.

A) densité électronique (Fo-Fc) à 2.74 Å de résolution (en vert) et Δ_{anom} à 5.0 Å de résolution (en violet), contours à 3σ , dans le site de fabrication du centre FeS.

B) Paire stéréo de la carte « omit » faite en enlevant du modèle le C β et le S γ de la Cys321.

Afin de localiser les métaux dans la carte de densité électronique, la contribution de la dispersion anormale a été analysée : alors qu’elle était extrêmement faible à une résolution de 2.74 Å, celle-ci n’est pas négligeable à une résolution de 5 Å. La position de ce pic de dispersion anormale coïncide avec celle de la carte de densité électronique avec des coefficients (Fo-Fc) (Figure II–16 A). En ajoutant et enlevant successivement des atomes de Fe ou S du modèle, 4 pics ont été identifiés dans des cartes du type (Fo-Fc) (numérotés de I à IV dans la Figure II–17 (a) à gauche). La position du pic I coïncide avec celle du pic de dispersion anormale, il a donc été assigné à un atome de fer. Le pic II étant à 3.6 Å du pic I, il peut être attribué à un atome de fer par comparaison à la structure d’Ala³⁸-IscU d’*Aquifex aeolicus* (Aa) (code PDB 2z7e) qui contient un centre [2Fe-2S] (Figure II–17 (a) et (b)). Le pic III se trouve entre les pics I et II qui correspondent à 2 atomes de Fe. Il a été, par conséquent, attribué à un atome de soufre pontant. Le pic correspondant au deuxième atome de soufre dans un centre [2Fe-2S] classique est absent de la densité électronique. S’il était présent, cet atome serait trop proche du S γ de la Cys102. En superposant la structure de notre complexe sur celle d’IscU d’Aa, on peut voir que le plan formé par les atomes I, II, III du premier est incliné par rapport à celui formé par le centre [2Fe-2S] du deuxième (Figure II–17 (b)). Le pic IV se trouve à 2.6 Å d’un des atomes de Fe et à 1.9 Å du pic III ; cette dernière distance pourrait correspondre à une liaison S-S.

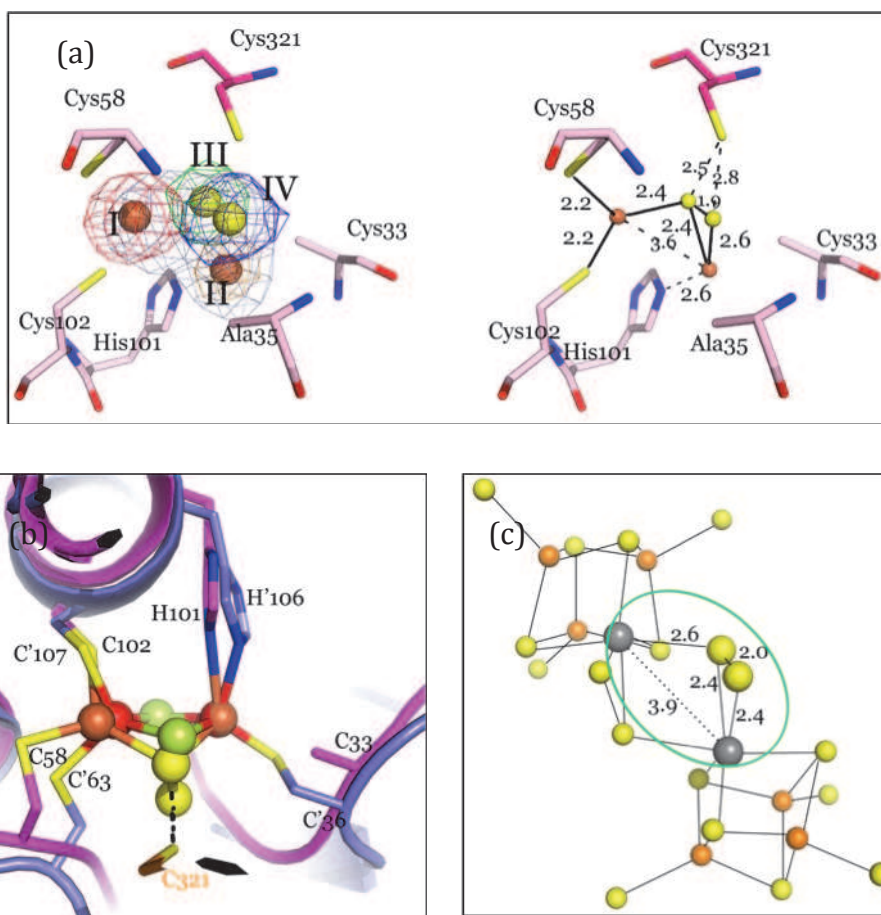


Figure II–17 Complexe IscSU exposé à l'air (AE).

(a) IscU est en rose, IscS en magenta. A gauche, le complexe avec son centre [2Fe-S-S] représenté par des sphères numérotées de I à IV avec la carte « omit » de chaque atome, ainsi que la carte « omit » globale du centre [2Fe-S-S] en gris (contours à 3σ). A droite les distances sont données en Å. L'atome de soufre de la Cys33 n'est pas visible dans la carte et, en conséquence, il n'a pas été inclus dans le modèle.

(b) Superposition de AE-A/IscSU sur AeIscU. AE-A/IscSU en rose et orange, son centre [2Fe-S-S] est en Fe = orange, S = jaune ; AeIscU en violet, son centre [2Fe-2S] est en Fe = rouge, S = vert.

(c) Centre Fer-Molybdène-Soufre, montrant une coordination analogue à celle trouvée dans le centre FeS de notre complexe AE ((Kovacs, Bashkin et al. 1985). Le centre Fe-Mo-S est représenté en sphères grises (Mo), jaunes (S) et oranges (Fe), avec la coordination du type $\mu_2\text{-}\eta^3$ mise en évidence par un cercle vert.

Le centre métallique semble être donc du type [2Fe-S-S] (Figure II–17 (a)) : un atome de fer est coordiné par les Cys58 et Cys102, le 2^{ème} atome de fer interagit avec l'His101, un atome de soufre (à proximité de la Cys321) est pontant entre les deux atomes de fer et forme une liaison S-S avec l'autre atome de S lui aussi à proximité de la Cys321. La coordination $\mu_2\text{-}\eta^3$ des ions fer au ligand persulfuré est similaire à celle observée pour un centre fer-molybdène-soufre dans un composé modèle du FeMoco de la nitrogénase (Kovacs, Bashkin et al. 1985) (Figure II–17 (c)).

Bien que le complexe semble stable à l'air (d'après nos expériences de spectroscopie UV-visible (II.3.1)) et (Raulfs, O'Carroll et al. 2008) la forme AE semble être deux électrons plus oxydée qu'un centre [2Fe-2S] standard. Dans le but de confirmer cela, nous avons cristallisé le complexe en condition anaérobie, en BAG, en présence de DTT et nous avons résolu sa structure.

II.3.5) Structure de IscSU traitée au DTT

La structure du complexe DT a été résolue à 2.53 Å de résolution par affinement restreint en utilisant la structure du complexe AE sans son agrégat FeS, le groupe d'espace et les paramètres de maille étant les mêmes (Tableau II-2). Le modèle a été corrigé avec le logiciel graphique COOT et l'affinement a été réalisé avec les logiciels AUTOBUSTER (Blanc, Roversi et al. 2004) et REFMAC5 (Murshudov, Vagin et al. 1997).

La structure de DT IscSU est différente de celle d'AE IscSU. Les deux molécules d'IscU sont visibles dans la carte de densité électronique (Figure II-18).



Figure II-18 Hétérotétramère de DT-Af(IscSU)₂.

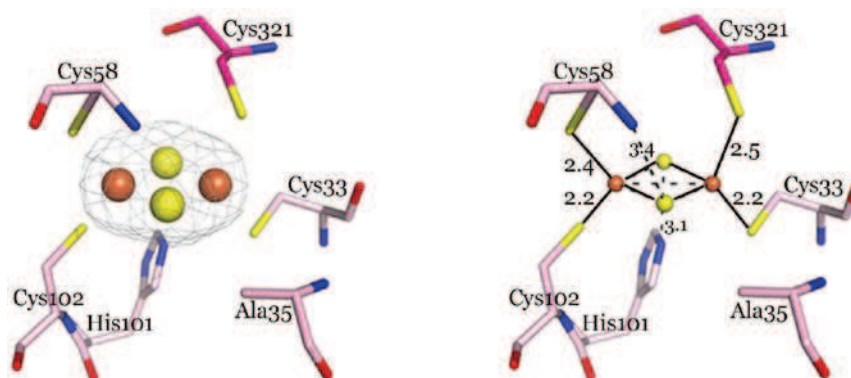
En rose et violet : IscU. En orange et vert : IscS. Les molécules de PLP sont représentées par des bâtons et sont indiquées par des flèches noires. Les cystéines conservées des protéines IscS sont représentées par des bâtons et sont indiquées par des flèches violettes.

La présence d'un centre $[2\text{Fe-2S}]$ dans le site d'assemblage est évidente par la forme de la densité électronique (Figure II–19). Un atome de fer est coordonné par les Cys321 et Cys33, un autre par les Cys58 et Cys102. Le centre $[2\text{Fe-2S}]$ est stabilisé par une interaction entre un des deux sulfures et le NH de la chaîne principale de la Cys58 et celle de l'autre sulfure avec un azote du cycle de l'His101. La boucle contenant la Cys321, ainsi que la boucle contenant la Cys33, sont plus ordonnées que dans le cas du complexe AE.

Les quantités de molécule de PLP (0.7 par IscS) et de centre FeS (1 par IscU comparé à 0.6 dans la forme AE) sont supérieures à celle observées en solution (environ 0.3 molécules de PLP et 0.3 centres FeS). La cristallisation sélectionnerait donc une forme de complexe riche en PLP et centre FeS.

Figure II–19 Complexe IscSU traité au DTT (DT).

IscU en rose, IscS en magenta. A gauche, le complexe avec son centre $[2\text{Fe-2S}]$ est représenté par des sphères et sa carte du type « omit » contournée à 3σ . A droite les distances sont données en Å.



II.3.6) Interface entre *IscS* et *IscU*

Si l'on superpose la forme AE sur la forme DT, on observe que le complexe est sensiblement le même (Figure II–20 (a)). Seul la position du résidu Cys321, et donc d'une portion de la boucle la contenant, diffère de moins d'1 Å.

Si l'on superpose *EcIscSU* sur la forme DT d'*AfIscSU* par leurs molécules A et C respectivement (IscU), on peut voir au niveau de l'interface entre IscS et IscU que la position de l'hélice contenant l'His350 d'*EcIscSU* équivalente à l'His343 d'*AfIscSU* diffère (Figure II–20 (b)). Pour *AfIscSU* la distance entre l'atome NE2 de l'His343 et le CO de la Cys33 est de 2.58Å ; pour *EcIscSU* la distance entre l'atome NE2 de l'His350 et le CO de la Cys33 est de 5.61Å. Dans le complexe IscSU d'*E.coli* la boucle contenant la cystéine n'est pas visible et les résidus visibles (325 à 327) avant cette boucle n'ont pas la même conformation que dans *Af*. Mis à part ces différences l'interface dans les deux complexes est similaire (Figure II–21).

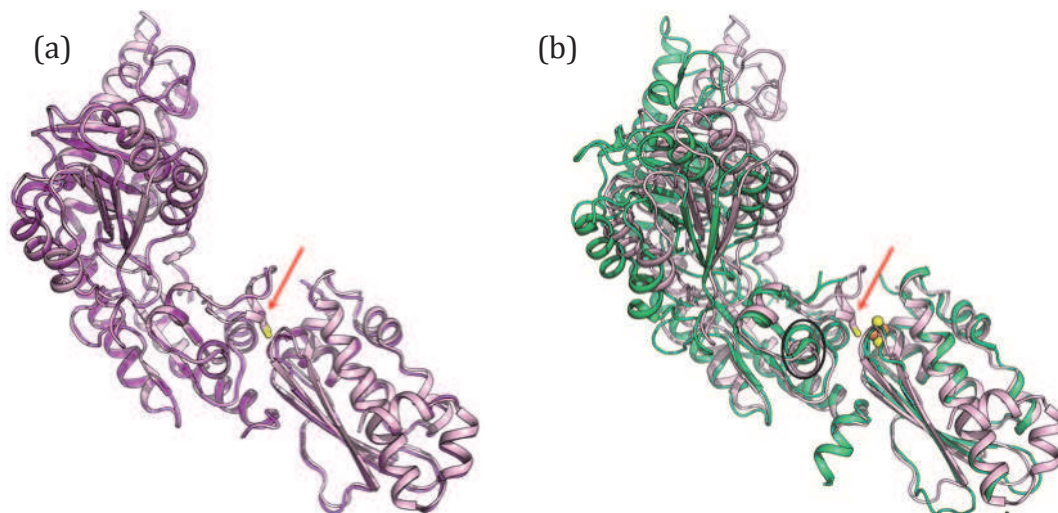


Figure II-20 Comparaison des formes AE/DT d'*AflscSU* et *EcIscSU*.

A. Superposition de la forme AE (violet) sur la forme DT (rose).

B. Superposition de la forme DT (rose) sur le complexe IscSU d'*Ec* (vert) (code PDB 3lvi). Le cercle noir indique la position différente de l'hélice contenant l'His350 d'*EcIscSU* équivalente à l'His343 d'*AflscSU*.

Les flèches rouges pointent sur la Cys321 d'*AflscSU*.

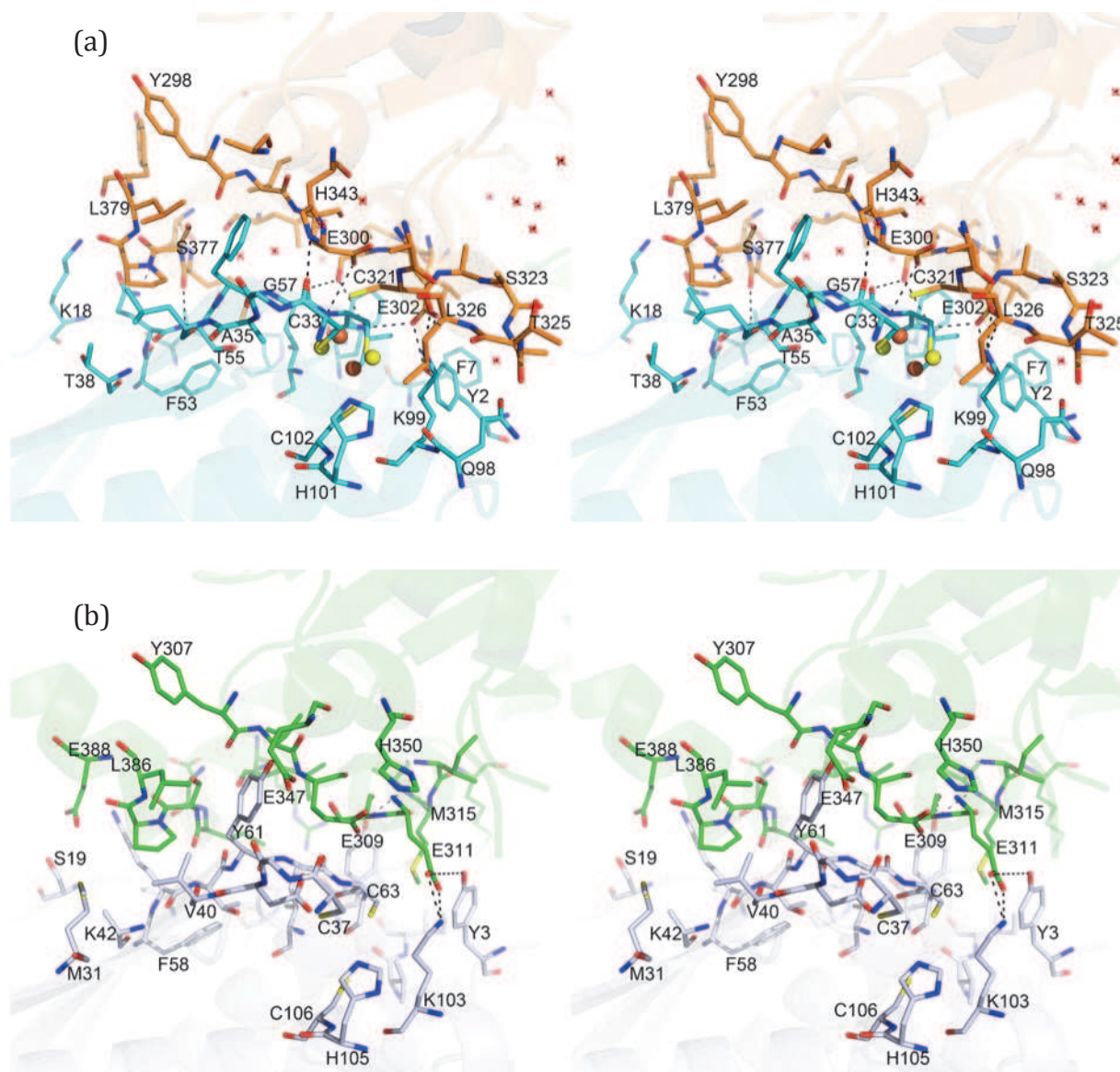


Figure II-21 Interface des complexes DT *A/IscSU* et *EcIscSU*.

Figures stéréo. (a) Complexe DT *A/IscSU*. IscS est représenté en orange et IscU en bleu. Les molécules d'eau sont représentées par des croix rouges. Le centre [2Fe-2S] est représenté par des sphères.

(b) Complexe *EcIscSU* (code PDB 3lv1, (Shi, Proteau et al. 2010)). IscS est représenté en vert et IscU en blanc.

La forme DT diffère de la forme AE par la liaison hydrogène entre le groupement carbonyle de la Cys33 d'IscU et le résidu His343 d'IscS. Cette interaction résulte de la boucle Cys33 ordonnée. La boucle contenant la Cys321 est elle aussi plus ordonnée (la Lys324 est visible). Des chaînes latérales de résidus sont visibles dans la forme DT alors qu'elles ne l'étaient que partiellement dans la forme AE (Tyr2, Lys99 par exemple) (remplacées par des molécules d'eau ou présence d'un petit pic positif dans la densité électronique (Fo-Fc)). La moins bonne résolution de la forme AE explique cette différence.

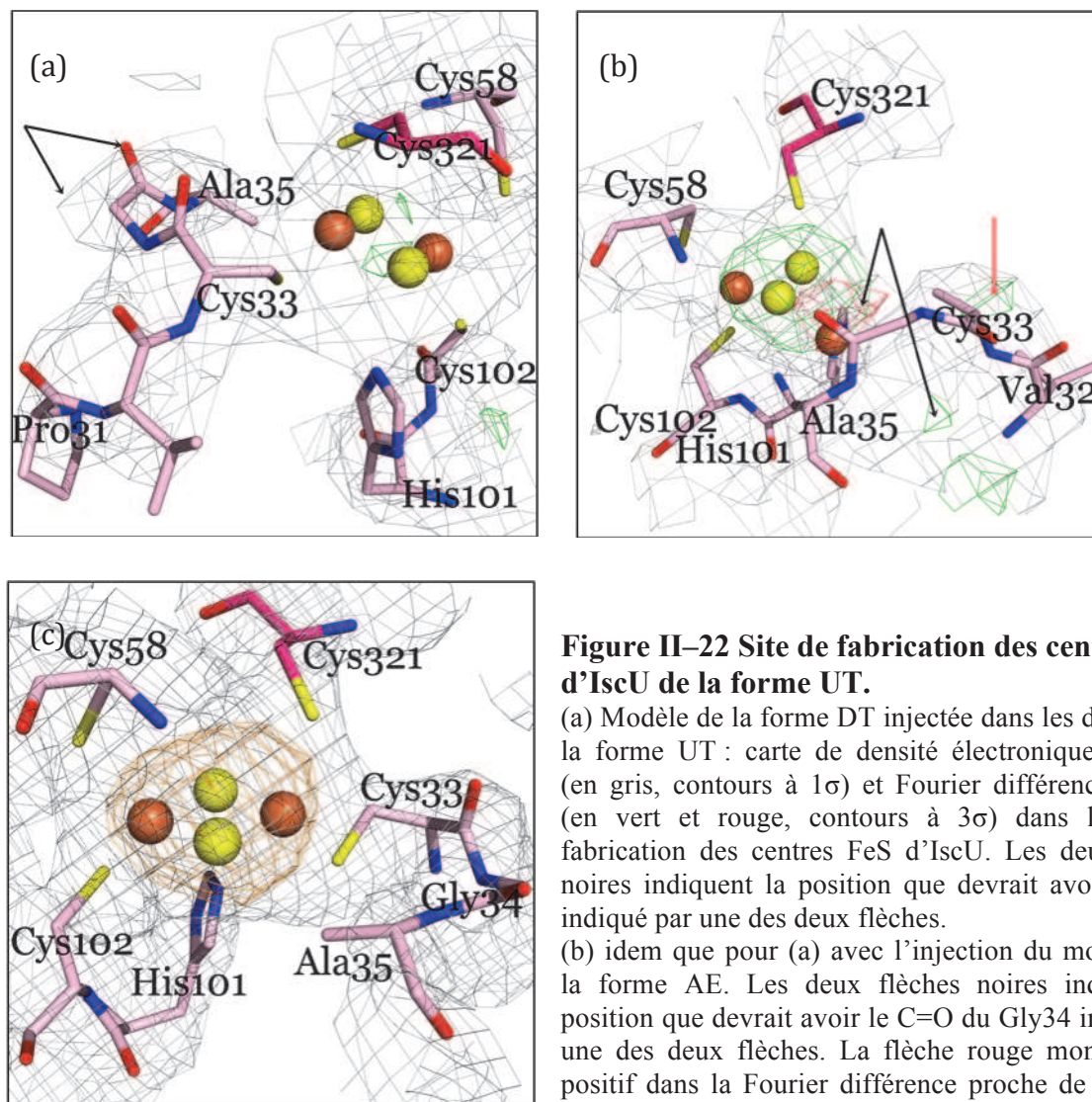
Les structures du complexe *A/IscSU* montrent que leur stabilité ne résulte pas de la formation d'un pont disulfure entre *IscS* et *IscU* comme décrit dans la référence (Kato, Mihara et al. 2002). Néanmoins en faisant une simple rotation autour de la liaison $C\alpha-C\beta$ de la Cys58 d'*IscU*, la Cys58 pourrait se lier à la Cys321.

Afin de savoir si la forme AE représente un intermédiaire d'assemblage des centres FeS (que l'on aurait donc réduit en ajoutant le réducteur DTT), ou si elle est le résultat de l'oxydation à l'air du complexe purifié en condition anaérobie pendant ou après la cristallisation sur la paillasse, ce complexe, sans ajout de DTT, a été cristallisé en BAG et sa structure a été résolue.

II.3.7) *Forme UT*

Le groupe d'espace dans le cas de la forme UT est le groupe $P2_1$ avec les paramètres de maille $a = 73.9 \text{ \AA}$, $b = 152.1 \text{ \AA}$, $c = 199.1 \text{ \AA}$, $\beta = 90.1^\circ$. Un des axes a doublé par rapport à la forme AE et DT de cristaux ce qui cause la perte de la symétrie orthorhombique. La structure du complexe de la forme UT a été résolue à 3.6 \AA de résolution par remplacement moléculaire, en utilisant la structure du dimère d'*IscS* (dans la structure de DT *IscSU* déterminée précédemment) en tant que modèle de recherche dans le programme PHASER (4 copies du dimère, selon le calcul des coefficients de Matthews (Matthews 1968)). La solution de remplacement moléculaire la plus probable trouvée par le programme a été retenue. N'ayant que des molécules d'*IscS* dans cette solution, les molécules d'*IscSU* dont la structure a été déterminée précédemment (forme AE ou DT) ont été superposées sur les molécules d'*IscS* afin d'obtenir les coordonnées des molécules d'*IscU*. Un affinement par corps rigide a été ensuite fait avec le logiciel REFMAC5 (Murshudov, Vagin et al. 1997) (Figure II–22).

Le modèle de la forme DT semble être celui correspondant le plus aux données de la forme UT, même s'il ne l'est peut être pas exactement le même, d'après la position du C=O de la Gly34.



II.3.8) Comparaison des formes DT/AE/UT d'AfIscSU

La forme AE résulte donc de l'oxydation à l'air du complexe purifié en condition anaérobie pendant ou après la cristallisation. Cette hypothèse est compatible avec le fait que la forme AE ne serait pas stable en solution au vu de son désordre (Figure II-17) au niveau de la coordination d'un des atomes de fer et la boucle contenant la Cys33.

Les spectre UV-vis sont restés inchangés après exposition à l'air du complexe IscSU (II.3.1)) parce que le passage de l'agrégat $[2\text{Fe}-2\text{S}]$ à la forme $[2\text{Fe}-\text{S}-\text{S}]$ préserve les bandes de transfert de charge de liaisons Fe-S.

Le passage de la forme DT ou UT à la forme AE est caractérisé par la formation d'un ligand persulfure : cette formation a requis un changement des positions de la Cys321 et d'un des atomes de fer ce qui a forcé la dissociation de celui-ci de la Cys33. En effet la mobilité de la boucle de la Cys33 dans le cristal devait être insuffisante pour se réajuster et rester lié au fer. De plus, ces changements ont amené du désordre dans le site d'assemblage du centre FeS puisque l'atome de soufre de la Cys33 n'est pas visible. L'His101 est ligand du centre [2Fe-S-S], probablement du fait de l'absence de l'Asp35. Cette ligation a déjà été observée pour *AvNifU* et *AvIscU* (Bandyopadhyay, Chandramouli et al. 2008) et dans la structure d'*AaAla*³⁸-IscU (code PDB 2z7e). Elle semble conférer de la stabilité à son centre [2Fe-2S] ; en effet, lié à l'histidine, celui-ci ne peut pas être transféré à des protéines cibles (Wu, Mansy et al. 2002). Le mutant Asp35Ala de la protéine *AvIscU* est aussi incapable de transférer un centre [4Fe-4S] à une apo-aconitase (Unciuleac, Chandramouli et al. 2007). L'acide aspartique conservé du site d'assemblage semble donc être impliqué dans le transfert du centre FeS (voir plus bas).

II.3.9) Calculs théoriques

Après la fabrication du centre FeS, IscS se dissocie de IscU, un ligand de ce centre FeS doit donc être remplacé puisque la Cys321 de IscS ne le lie plus. Le ligand de remplacement le plus probable est l'Asp35 (voir plus haut). En mutant l'Asp35 en alanine, nous avons obtenu un complexe (IscS-Ala³⁵-IscU)₂ contenant un centre [2Fe-2S]. Il semblerait donc que cet acide aspartique ne soit pas impliqué dans l'assemblage du centre. Si tel est le cas, il devrait y avoir assez de place dans notre structure pour le -COOH de cet acide aspartique.

Des calculs théoriques avec la suite Schrödinger (Schrödinger 2011) sur la forme du complexe DT ont été faits (voir Matériel et Méthodes). Les molécules A (IscS) et C (IscU) de la structure de la forme DT ont été extraites, et l'alanine 35 a été mutée, *in silico*, par un acide aspartique dans la molécule C. La géométrie du modèle a été optimisée en utilisant un potentiel hybride de mécanique quantique / mécanique moléculaire (MQ/MM) (Figure II-24 A). Le centre [2Fe-2S] et ses ligands protéiques sont décrits par la mécanique quantique alors que le reste du modèle est décrit par la mécanique moléculaire. La structure optimisée du modèle muté Ala35Asp montre très peu de changements structuraux avec la structure cristallographique du mutant Ala35 (Figure II-23). Ceci suggère que la coordination du centre [2Fe-2S] observée dans la forme DT n'est pas un artéfact de la mutation de l'acide aspartique en alanine. On peut donc conclure que la mutation n'a pas altéré la conformation de la

protéine puisqu'il reste assez de place pour l'acide aspartique. L'acide aspartique n'interviendrait donc pas dans l'assemblage du centre.

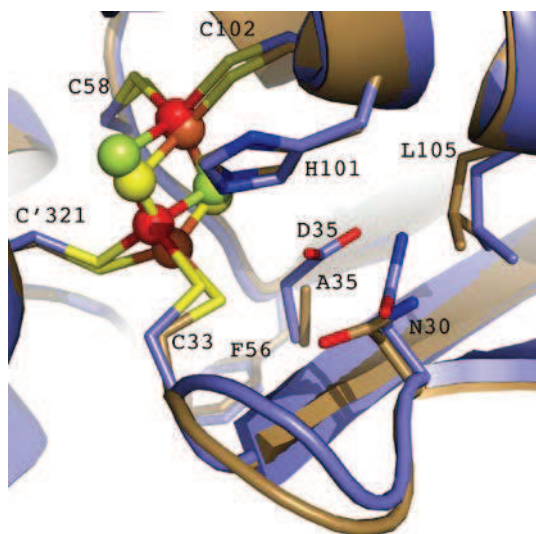


Figure II–23 Superposition de la forme DT sur le modèle MQ/MM optimisé où l'Ala35 a été substituée par l'Asp.

En brun le modèle cristallographique de la forme DT (mutant Ala35) et en bleu le modèle théorique de la forme Asp35.

Lors de la dissociation de IscS à IscU, le ligand Cys321 du centre doit être remplacé. L'Asp35 aurait donc plutôt un rôle dans le remplacement de ce ligand, puisque l'on a en plus isolé l'intermédiaire (IscS-Ala³⁵-IscU)₂ contenant un centre [2Fe-2S] grâce à la mutation. Afin de vérifier si la coordination du centre par la Cys33, l'Asp35, les cystéines 58 et 102 était possible, la boucle Cys33 (résidus 30 à 35), du modèle obtenu précédemment, adoptant la conformation de plus basse énergie a été cherchée par échantillonnage de boucle tout en restreignant les atomes S γ de la Cys33 et un des O δ de l'Asp35 à proximité de l'ion métallique. Cette contrainte a été imposée pour trouver des conformations de boucle qui soient pertinentes ; si aucune conformation avec la Cys33 et l'Asp35 n'avait été favorable, nous n'aurions pas obtenu une structure valide. La géométrie du modèle a alors été optimisée par calculs de MQ/MM. Le modèle final possède un centre [2Fe-2S] classique, coordonné par les cystéines 58, 102 et 33 et l'Asp35 (Figure II–24 B). L'Asp35 pourrait donc remplacer la Cys321 d'IscS lors de la dissociation de IscS à IscU. Le remplacement du ligand Cys321 implique une faible translation du centre [2Fe-2S] avec les Cys33 et l'Asp35 liant alors l'atome de Fe (Figure II–24).

La comparaison des deux structures des modèles de MQ/MM suggère que lors de la dissociation du complexe IscSU, le remplacement du ligand Cys321 implique un réarrangement de la Cys33 et de l'Asp35. Ce changement devrait permettre d'éviter la labilisation de l'agrégat [2Fe-2S] qui aurait lieu lors de la dissociation d'IscS à IscU s'il n'y

avait pas de substitution de ligands, du fait d'une exposition excessive au solvant (Figure II–25).

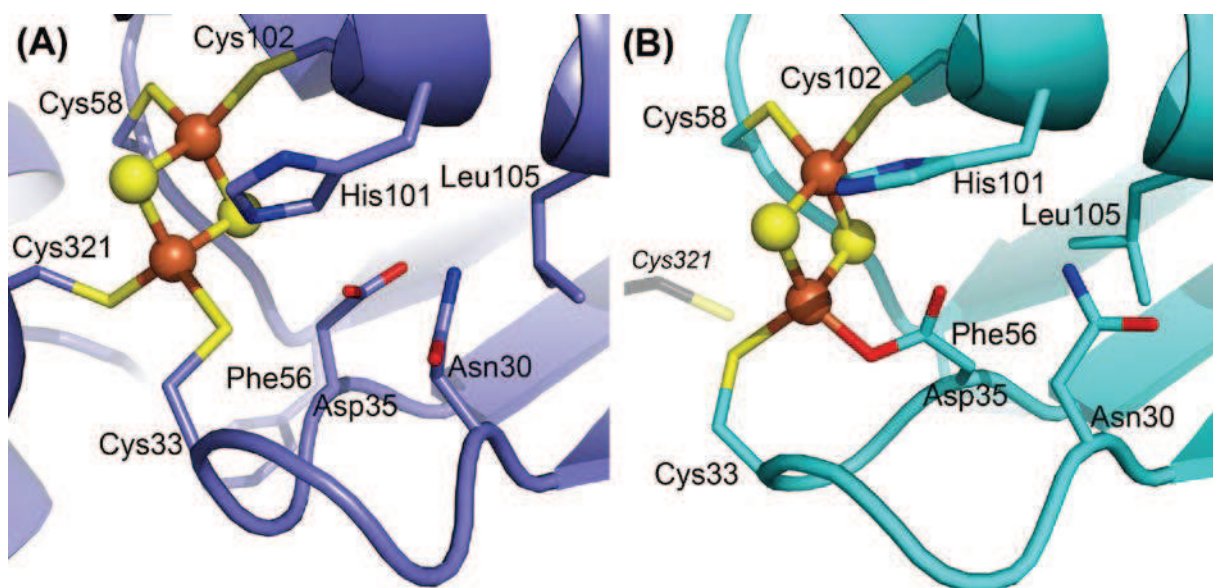
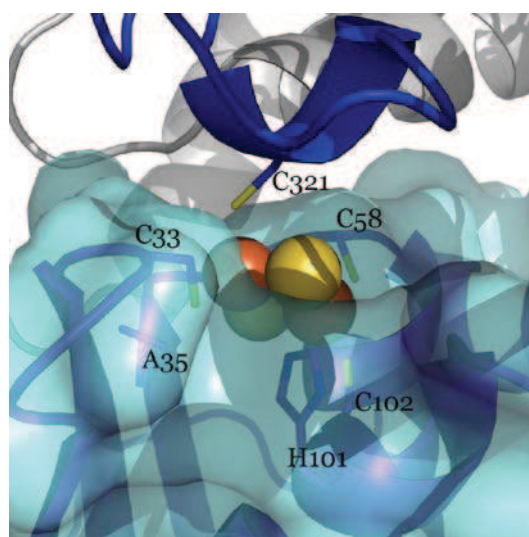


Figure II–24 Modèles MQ/MM.

A. Modèle MQ/MM *Af*(IscS-Asp³⁵-IscU)₂ dérivé de la forme DT. (B) Modèle MQ/MM *Af*(Asp³⁵-IscU) où l'échantillonnage de la boucle 30-35 a été fait.

Figure II–25 Interface d'IscS/IscU de la forme DT.

En bleu foncé, IscS ; en bleu clair IscU.



La boucle flexible de la Cys33 semble alors agir comme une « pince» sur l'atome de fer le plus exposé, avec ses résidus Cys33 et Asp35, séparés par une glycine, pendant la dissociation du complexe IscSU. Le même type de « pince » peut être observé dans la structure d'IscU de

Streptococcus pyogenes (code PDB 1su0) où la Cys40 et l'Asp42 équivalents liant un ion Zn^{2+} , mais dans ce cas la pince n'a probablement pas de fonction physiologique (Figure II–26). Selon sa stabilité, IscU peut alors s'associer à d'autres protéines comme par exemple en formant un dimère avec une apo-IscU (Agar, Krebs et al. 2000). Le centre $[2\text{Fe}-2\text{S}]$ sera alors soit directement transféré à une apoprotéine cible, soit deux $[2\text{Fe}-2\text{S}]$ -IscU vont s'assembler pour former un centre $[4\text{Fe}-4\text{S}]$. Le centre FeS peut être stocké transitoirement sur des ATCs (voir II.2.2.3) de l'introduction) qui ont alors la capacité de transférer ce centre à des apoprotéines cibles.

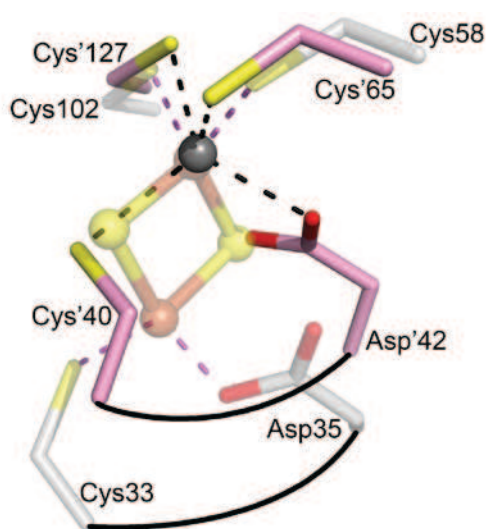


Figure II–26 « Pince » Cys33-Asp35.

En bleu clair la « pince » pour le modèle MQ/MM dérivé de la forme AE et en rose celle de *Streptococcus pyogenes* IscU (code PDB 1su0).

II.3.10) Spectres RPE et Raman

Afin d'étudier l'état redox et la composition du complexe IscSU (dans le tampon 50 mM Tris pH 8 600 mM NaCl) traité ou non au DTT, des échantillons préparés en condition anaérobie en BAG ont été analysés par spectroscopie RPE (voir Matériel et Méthodes). Du DTT en concentration finale 5mM a été ajouté à un des échantillons, pour un temps d'incubation d'une nuit à 4°C. L'échantillon a ensuite été dialysé et concentré afin d'éliminer le DTT et d'atteindre une concentration de protéine d'environ 0.5 mM. L'échantillon sans traitement au DTT a lui seulement été concentré à environ 0.5 mM.

Les deux échantillons se sont avérés être silencieux à la spectroscopie RPE. Le centre $[2\text{Fe}-2\text{S}]$ serait donc composé de deux atomes de Fe^{2+} ou de deux atomes de Fe^{3+} , avec une densité de spin nulle due à un couplage antiferromagnétique. Pour assembler un agrégat $[2\text{Fe}-2\text{S}]$ la Cys321 doit donner, successivement, 2 atomes S^0 à IscU. Donc, un total de 4 électrons

est nécessaire pour générer leurs ligands sulfure. La première source de réduction pourrait provenir de l'oxydation des deux ions ferreux (Krebs, Agar et al. 2001) en ions ferriques. On peut donc supposer qu'il y a deux ions Fe^{3+} dans le centre $[\text{2Fe-2S}]$ de notre complexe A/IscSU . La seconde source d'électrons ferait intervenir un donneur extérieur (discuté plus tard).

Une expérience de spectroscopie Raman sur les cristaux dans leur solution de cristallisation a été réalisée. Les cristaux de la protéine prétraitée à 5 mM DTT (Figure II–27) présentent les bandes caractéristiques d'un centre $[\text{2Fe-2S}]$ (Tableau II–3). Les cristaux de la protéine non traitée au DTT présentent les mêmes bandes que précédemment, ce qui est cohérent avec nos résultats cristallographiques. Les échantillons avant et après traitement au DTT semblent donc être les mêmes comme vu dans les cristaux. Les cristaux de la forme AE n'ont malheureusement pas pu être testés car nous n'arrivons pas à les reproduire.

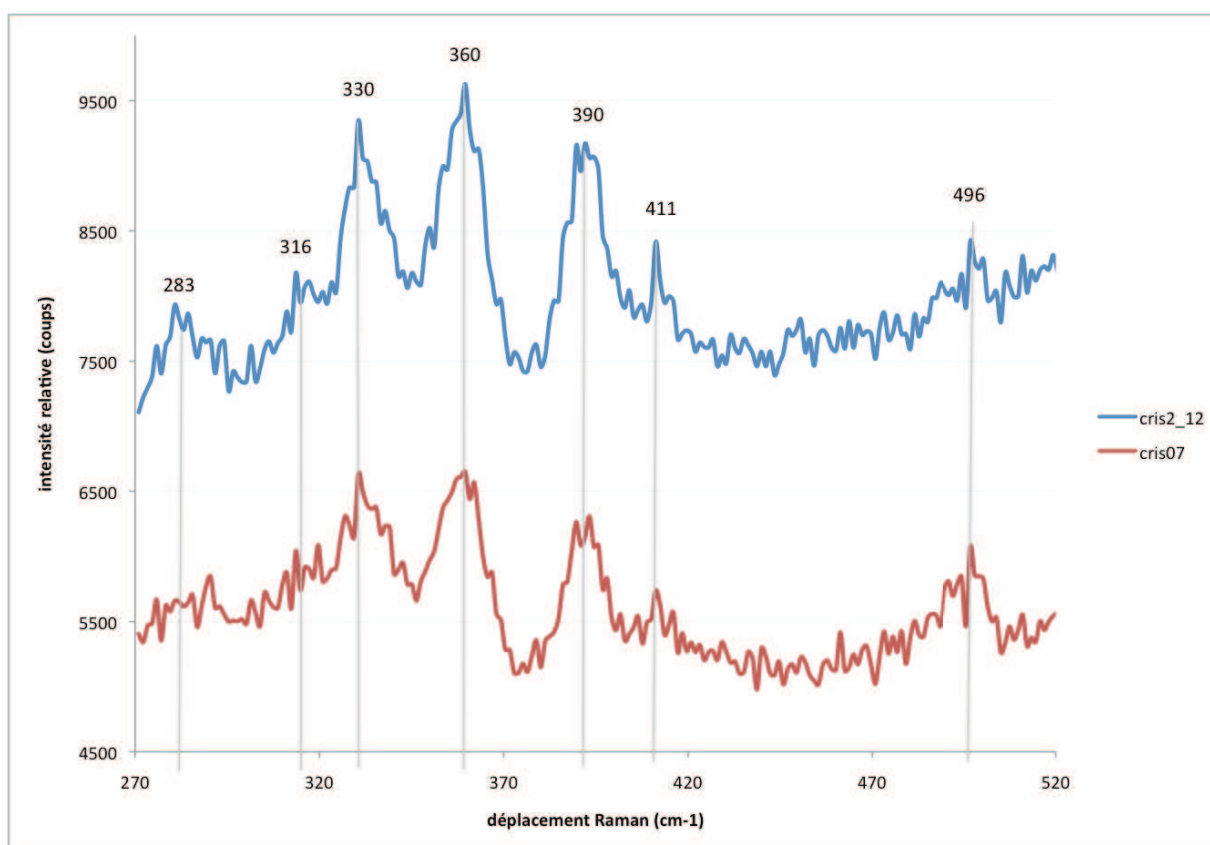


Figure II–27 Spectres Raman.

Expérience réalisée à T_{amb} , 514nm. cris2-12 : cristaux de la protéine prétraitée à 5 mM DTT (DT) ; cris07 : cristaux de la protéine sans traitement au DTT (UT). La bande à environ 500 cm^{-1} pourrait correspondre à une vibration S-S mais elle n'a pas encore été vérifiée par, par exemple, substitution isotopique.

	Déplacements Raman (cm ⁻¹)									
<i>A</i> /IscSU			283	316	330	360	390	411		500
<i>Av</i> IscU (Agar, Krebs et al. 2000)			296		328	356	370	407	423	/
(Han, Czernuszewicz et al. 1989) Adrenodoxine, Ado Ferrédoxine, Fd		282	291	317	329 329	341 339	349 357, 367	393 395	421 426	/ /

Tableau II–3 Bandes Raman d’un centre [2Fe-2S].

Données obtenues expérimentalement et déjà publiées. Les « / » indiquent que les spectres Raman ne couvrent pas ces valeurs.

II.3.11) Proposition d’un mécanisme de formation des centres FeS

En superposant la structure de notre forme DT, où la cystéine active se trouve à proximité du site d’assemblage des centres FeS d’IscU, sur la structure de *Synechocystis* SufS (code PDB 1t3i), pour laquelle la cystéine active fait face à la molécule de PLP, on peut voir que cette cystéine fait un mouvement d’environ 14 Å du PLP vers le site d’assemblage (Figure II–28).



Figure II–28 Différentes conformations de la boucle contenant la cystéine active.

En rose la forme DT du *A*/IscS et en rose clair SufS de *Synechocystis* (code PDB 1t3i).

Plusieurs observations amènent à penser que la Cys321 peut donner successivement des atomes de soufre à IscU sans dissociation du complexe : 1) le complexe (IscS-Ala³⁵-IscU)₂ est stable (Raulfs, O’Carroll et al. 2008) et 2) la formation du complexe IscS-IscU ne nécessite pas d’avoir une boucle fixe comme vu pour le complexe d’*Ec* (Shi, Proteau et al. 2010) (Figure II–20) ; et 3) il y a un large tunnel connectant le site d’assemblage des centres FeS d’IscU et le site actif d’IscS permettant la diffusion de la L-cystéine libre et les changements conformationnels de boucle (Figure II–29). La boucle pourrait donc osciller entre une

conformation où la cystéine se trouve dans le site actif de IscS et une autre où elle se trouve dans le site d'assemblage du centre FeS d'IscU.

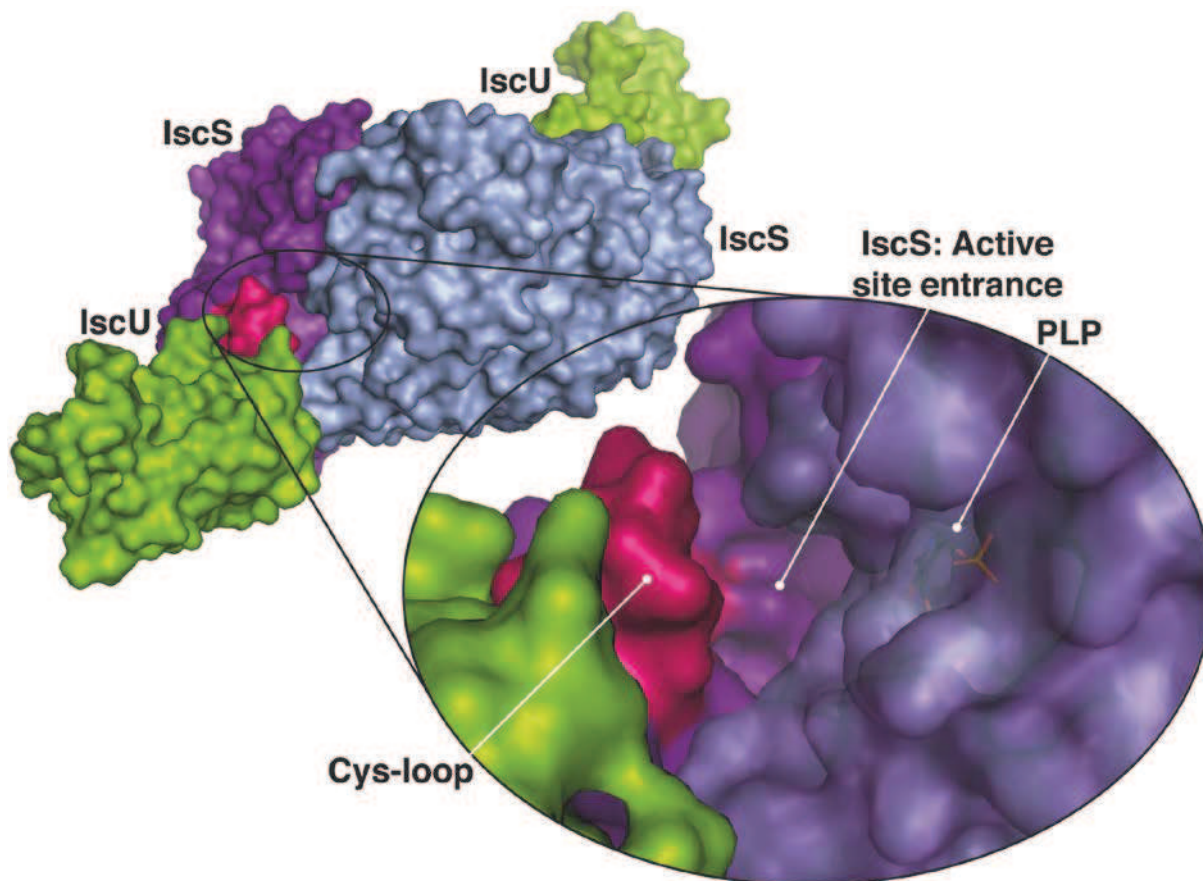


Figure II–29 Représentation de la surface du complexe IscSU.

La surface du complexe IscSU est représentée. Dans l'encadré, une vue rapprochée sur le site de fixation du PLP et la boucle de la Cys321 (Cys-loop), la surface de la molécule IscS cachant le PLP est représentée en semi-transparent.

La première étape de fabrication des centres FeS serait la formation du complexe (IscSU)₂, état représenté par la structure du complexe d'*E.coli* (Figure II–20).

Il y a eu une controverse concernant l'ordre d'arrivée des atomes de fer et soufre dans la fabrication du centre FeS. Concernant l'hypothèse d'arrivée première du soufre puis du fer, il a été observé que les 3 résidus cystéine d'IscU peuvent être poly-sulfurés en l'absence de fer (Urbina, Silberg et al. 2001), (Smith, Agar et al. 2001), ce qui donne la possibilité de transfert intra ou inter moléculaire de soufre entre les cystéines conservées d'IscS et IscU durant le processus d'assemblage du centre FeS. Néanmoins la persulfuration de *Thermotoga maritima*

(*Tm*) apo-IscU a amené à un intermédiaire bloqué qui ne liait pas le fer (Nuth, Yoon et al. 2002).

La deuxième hypothèse d'arrivée première du fer puis du soufre a été étudiée en observant la liaison de fer aux formes apo de *Av*NifU et *Av*NifU-1 (NifU possède 3 domaines, NifU-1 correspond au domaine du type IscU) (Agar, Yuvaniyama et al. 2000) et *Tm*IscU (Nuth, Yoon et al. 2002). Le fer se lie à ces protéines mais à des températures plus basses que les températures physiologiques ($T=2^{\circ}\text{C}$ pour *Av*, organisme mésophile ; $T=27^{\circ}\text{C}$ pour *Tm*, organisme hyperthermophile).

Enfin une troisième hypothèse d'arrivée simultanée du fer et du soufre peut être proposée. Il est supposé que dans la mitochondrie le premier transfert de fer de la frataxine à Isu1 (l'orthologue eucaryote de IscU (levure)) a lieu sur un site à proximité de celui de fabrication des centres FeS (Cook, Kondapalli et al. 2010). Ceci laisse penser à un mécanisme où l'arrivée du fer au site d'assemblage des centres FeS pourrait être provoqué par l'arrivée de la Cys321 dans ce site.

La forme AE du complexe IscSU suggère que le site d'assemblage des centres FeS peut contenir un ligand pontant persulfido libre ou sous la forme Cys321-S-S⁻ avec une coordination du type μ_2 avec les deux atomes de fer. Pour assembler un agrégat [2Fe-2S] la Cys321 doit donner, successivement, 2 atomes S⁰ à IscU. Comme vu précédemment, la première source de réduction proviendrait de l'oxydation des deux ions ferreux (Krebs, Agar et al. 2001). La seconde aurait besoin d'un donneur extérieur. La libération du second atome de soufre peut impliquer le transfert non-rédox du S⁰ de la Cys321 persulfurée vers le S²⁻ déjà présent dans le site d'assemblage pour former le ligand persulfido des atomes de fer, ou le premier ligand sulfido pourrait être déplacé de l'autre côté des fers.

Le mécanisme de fabrication des centres FeS suggéré par nos résultats et la littérature est le suivant : 1) formation d'un complexe apo IscS-IscU; 2) persulfuration de la Cys321 d'IscS; 3) arrivée simultanée de la cystéine vers le site de fabrication des centres FeS d'IscU et des atomes de fer; 4) réduction du S⁰ en S²⁻ dont les électrons proviennent de l'oxydation des 2 ions Fe²⁺ en Fe³⁺; 5) seconde persulfuration de la Cys321 sans dissociation du complexe ; 6) liaison de la cystéine persulfurée et probable transfert du S⁰ au ligand sulfure présent dans le site de fabrication du centre FeS, formant un centre [2Fe-S-S] ; 7) réduction du centre FeS par 2 e⁻ fournis par un donneur extérieur pour former un centre [2Fe-2S] classique coordonné par les 3 Cys d'IscU et la Cys321 (représenté dans la structure DT-*A*/IscSU (Figure II-20)); 8) dissociation du complexe IscSU impliquant le réarrangement de la coordination du Fe avec le remplacement de la Cys321 par l'Asp35.

II.4) Conclusion et perspectives

La surexpression des protéines Ala³⁵-IscU et IscS d' *Archaeoglobus fulgidus* permet d'isoler un complexe stable $Af(IscS-Ala^{35}-IscU)_2$ contenant un centre FeS. Un complexe similaire avait été isolé pour *Azotobacter vinelandii* (Raulfs, O'Carroll et al. 2008). Lors de la résolution de la structure de la protéine IscS d' *A. fulgidus*, la boucle contenant la cystéine conservée d'IscS (Cys321) n'a pas pu être observée. Elle n'avait également pas pu être visualisée dans les structures des différentes IscS publiées précédemment. L'absence de lysine formant une base de Schiff avec le PLP différencie cette cystéine désulfurase des autres. Du fait de sa faible activité, il est possible qu'elle possède un partenaire protéique *in vivo* qui l'augmenterait. Le complexe $Af(IscS-Ala^{35}-IscU)_2$ exprimé dans *E. coli* étant chargé en centres FeS, ce partenaire aurait alors un homologue dans cet organisme.

Trois formes de cristaux de ce complexe ont été obtenues dans les conditions suivantes:

- après exposition à l'air d'un échantillon purifié en condition anaérobie (AE),
- après traitement au DTT de l'échantillon purifié en condition anaérobie (DT),
- à partir de l'échantillon purifié en condition anaérobie (UT).

Pour chacune de ces formes cristallines, la résolution de la structure du complexe a permis l'observation de la boucle contenant la Cys321 d'IscS, proche du site d'assemblage des agrégats FeS dans IscU. La forme AE contient un centre [2Fe-S-S] dans ce site. L'agrégat est coordonné par l'His101 et par Cys58 et Cys102. Le thiolate de la Cys321 est proche (à 2.5 Å) d'un des atomes de soufre du ligand sulfido. La boucle contenant la Cys33 est partiellement désordonnée, le S γ de ce résidu n'est pas visible dans la structure. L'Ala35 est dans l'environnement de ce centre FeS.

Dans le cas de la forme DT, la Cys321 est ligand d'un des atomes de fer d'un centre [2Fe-2S]. Les autres ligands de l'agrégat sont les cystéines conservées de IscU, Cys33, Cys58 et Cys102. La forme UT étant iso-structurale à la forme DT, la forme AE est donc due à l'oxydation à l'air du complexe purifié en condition anaérobie pendant ou après la cristallisation. La mutation Asp→Ala permettant l'obtention d'un complexe IscSU stable, il semblerait que l'acide aspartique conservé d'IscU soit impliqué dans la dissociation du complexe. Des calculs théoriques confirment la possibilité du remplacement du ligand Cys321 par l'Asp35 lors de la dissociation du complexe IscS-IscU.

La proposition d'un mécanisme de formation des centres FeS a pu être appuyée par ces observations et les données de la littérature.

Les perspectives suite à l'obtention de ces résultats sont :

- de faire des nouveaux essais de reproduction des cristaux de la forme AE, afin d'obtenir le spectre Raman de cette forme,
- obtenir d'autres intermédiaires d'assemblage d'un centre [2Fe-2S] avec le mutant D35A d'*AflscU* ou la protéine native:
 - faire une étude cristallographique en ajoutant du soufre seul, soufre + fer, soufre + fer + réducteur au complexe IscS-IscU cristallisé,
 - faire la même étude avec les protéines d'*E. coli*,
- comprendre dans quelles conditions *AflscS* est active:
 - a-t-elle réellement une activité cystéine désulfurase? Pour cela une mutation de la cystéine en sérine pourrait être faite, et voir si l'on obtient toujours le complexe chargé en centre [2Fe-2S],
 - une recherche d'un éventuel partenaire *in vivo* pourrait être faite. Dans ce contexte les gènes adjacents et notamment DsrE pourraient être étudiés.

Bibliographie

- Agar, J. N., C. Krebs, et al. (2000). "IscU as a scaffold for iron-sulfur cluster biosynthesis: Sequential assembly of 2Fe-2S and 4Fe-4S clusters in IscU." *Biochemistry* **39**(27): 7856-7862.
- Agar, J. N., P. Yuvaniyama, et al. (2000). "Modular organization and identification of a mononuclear iron-binding site within the NifU protein." *J Biol Inorg Chem* **5**(2): 167-177.
- Agar, J. N., L. M. Zheng, et al. (2000). "Role of the IscU protein in iron-sulfur cluster biosynthesis: IscS-mediated assembly of a [Fe₂S₂] cluster in IscU." *J Am Chem Soc* **122**: 2136-2137.
- Alliegro, M. C. (2000). "Effects of dithiothreitol on protein activity unrelated to thiol-disulfide exchange: for consideration in the analysis of protein function with Cleland's reagent." *Anal Biochem* **282**(1): 102-106.
- Altschul, S. F., W. Gish, et al. (1990). "Basic local alignment search tool." *J Mol Biol* **215**(3): 403-410.
- Arragain, S., R. Garcia-Serres, et al. (2010). "Post-translational modification of ribosomal proteins : structural and functional characterization of RimO from *Thermotoga maritima*, a radical S-Adenosylmethionine methylthiotransferase " *J Biol Chem* **285**(8): 5792-5801.
- Bandyopadhyay, S., K. Chandramouli, et al. (2008). "Iron-sulfur cluster biosynthesis." *Biochem Soc T* **36**: 1112-1119.
- Behshad, E. and J. M. Bollinger (2009). "Kinetic analysis of cysteine desulfurase CD0387 from *Synechocystis* sp PCC 6803: formation of the persulfide intermediate." *Biochemistry* **48**(50): 12014-12023.
- Behshad, E., S. E. Parkin, et al. (2004). "Mechanism of cysteine desulfurase Slr0387 from *Synechocystis* sp PCC 6803: Kinetic analysis of cleavage of the persulfide intermediate by chemical reductants." *Biochemistry* **43**(38): 12220-12226.
- Beinert, H., M. C. Kennedy, et al. (1996). "Aconitase as iron-sulfur protein, enzyme, and iron-regulatory protein." *Chem Rev* **96**(7): 2335-2373.
- Berkovitch, F., Y. Nicolet, et al. (2004). "Crystal structure of biotin synthase, an S-adenosylmethionine-dependent radical enzyme." *Science* **303**(5654): 76-79.
- Blanc, E., P. Roversi, et al. (2004). "Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT." *Acta Crystallogr D* **60**: 2210-2221.
- Boal, A. K., T. L. Grove, et al. (2011). "Structural basis for methyl transfer by a radical SAM enzyme." *Science* **332**(6033): 1089-1092.
- Boeck, A., P. W. King, et al. (2006). Maturation of hydrogenases. *Adv Microb Physiol*. R. K. Poole. **51**: 1-71.
- Brunger, A. T. (1992). "FREE R-VALUE - A NOVEL STATISTICAL QUANTITY FOR ASSESSING THE ACCURACY OF CRYSTAL-STRUCTURES." *Nature* **355**(6359): 472-475.
- Burmeister, W. P. (2000). "Structural changes in a cryo-cooled protein crystal owing to radiation damage." *Acta Crystallogr D* **56**: 328-341.
- Buurman, G., S. Shima, et al. (2000). "The metal-free hydrogenase from methanogenic archaea: evidence for a bound cofactor." *FEBS Lett* **485**(2-3): 200-204.
- Challand, M. R., F. T. Martins, et al. (2010). "Catalytic activity of the anaerobic tyrosine lyase required for thiamine biosynthesis in *Escherichia coli*." *J Biol Chem* **285**(8): 5240-5248.
- Chandramouli, K. and M. K. Johnson (2006). "HscA and HscB stimulate 2Fe-2S cluster transfer from IscU to apoferredoxin in an ATP-dependent reaction." *Biochemistry* **45**(37): 11087-11095.
- Chandramouli, K., M.-C. Unciuleac, et al. (2007). "Formation and properties of 4Fe-4S clusters on the IscU scaffold protein." *Biochemistry* **46**(23): 6804-6811.
- Chatterjee, A., Y. Li, et al. (2008). "Reconstitution of ThiC in thiamine pyrimidine biosynthesis expands the radical SAM superfamily." *Nat Chem Biol* **4**(12): 758-765.
- Chen, J. S. and L. E. Mortenson (1977). "Inhibition of Methylene-Blue formation during determination of acid-labile sulfide of iron-sulfur protein samples containing dithionite." *Anal Biochem* **79**: 157-165.
- Cook, J. D., K. C. Kondapalli, et al. (2010). "Molecular details of the yeast frataxin-Isul interaction during mitochondrial Fe-S cluster assembly." *Biochemistry* **49**(40): 8756-8765.
- Cupp-Vickery, J. R., H. Urbina, et al. (2003). "Crystal structure of IscS, a cysteine desulfurase from *Escherichia coli*." *J Mol Biol* **330**: 1049-1059.
- Dahl, C., A. Schulte, et al. (2008). "Structural and molecular genetic insight into a widespread sulfur oxidation pathway." *J Mol Biol* **384**(5): 1287-1300.
- Darnault, C., A. Volbeda, et al. (2003). "Ni-Zn- Fe₄-S₄ and Ni-Ni- Fe₄-S₄ clusters in closed and open subunits of acetyl-CoA synthase/carbon monoxide dehydrogenase." *Nat Struct Biol* **10**(4): 271-279.
- Demick, J. M. and W. N. Lanzilotta (2011). "Radical SAM activation of the B(12)-independent glycerol dehydratase results in formation of 5'-Deoxy-5'-(methylthio)adenosine and not 5'-Deoxyadenosine." *Biochemistry* **50**(4): 440-442.

- Drenth, J. (2007). Principles of Protein X-Ray Crystallography, Springer.
- Driesener, R. C., M. R. Challand, et al. (2010). "FeFe-Hydrogenase cyanide ligands derived from S-Adenosylmethionine-dependent cleavage of tyrosine." Angew Chem Int Edit **49**(9): 1687-1690.
- Ducruix, A. and R. Giegé (1999). Crystallization of Nucleic Acids and Proteins, Oxford University Press.
- Einsle, O., F. A. Tezcan, et al. (2002). "Nitrogenase MoFe-protein at 1.16 angstrom resolution: A central ligand in the FeMo-cofactor." Science **297**(5587): 1696-1700.
- Elliott, J. I. and J. M. Brewer (1978). "Inactivation of yeast enolase by 2,3-butanedione." Arch Biochem Biophys **190**(1): 351-357.
- Emsley, P., B. Lohkamp, et al. (2010). "Features and development of Coot." Acta Crystallogr D **66**: 486-501.
- Evans, P. and A. McCoy (2008). "An introduction to molecular replacement." Acta Crystallogr D **64**: 1-10.
- Farrar, C. E. and J. T. Jarrett (2009). "Protein residues that control the reaction trajectory in S-Adenosylmethionine radical enzymes: mutagenesis of asparagine 153 and aspartate 155 in *Escherichia coli* Biotin Synthase." Biochemistry **48**(11): 2448-2458.
- Fiser, A. and A. Sali (2003). MODELLER: Generation and refinement of homology-based protein structure models. Method Enzymol. C. W. S. R. M. Carter. **374**: 461-+.
- Fish, W. W. (1988). "Rapid colorimetric micromethod for the quantitation of complexed iron in biological samples." Method Enzymol **158**: 357-364.
- Flint, D. H. (1996). "*Escherichia coli* contains a protein that is homologous in function and N-terminal sequence to the protein encoded by the nifS gene of *Azotobacter vinelandii* and that can participate in the synthesis of the Fe-S cluster of dihydroxy-acid dehydratase." J Biol Chem **271**(27): 16068-16074.
- Fontecave, M. (2006). "Iron-sulfur clusters: ever-expanding roles." Nat Chem Biol **2**(4): 171-174.
- Fontecave, M. and S. Ollagnier-De-Choudens (2008). "Iron-sulfur cluster biosynthesis in bacteria: Mechanisms of cluster assembly and transfer." Arch Biochem Biophys **474**(2): 226-237.
- Fontecilla-Camps, J. C., P. Amara, et al. (2009). "Structure-function relationships of anaerobic gas-processing metalloenzymes." Nature **460**(7257): 814-822.
- Fontecilla-Camps, J. C., A. Volbeda, et al. (2007). "Structure/function relationships of NiFe - and FeFe - hydrogenases." Chem Rev **107**(10): 4273-4303.
- Frey, P. A., A. D. Hegeman, et al. (2008). "The radical SAM superfamily." Crit Rev Biochem Mol **43**(1): 63-88.
- Fu, W. G., R. F. Jack, et al. (1994). "NifU gene-product from *Azotobacter vinelandii* is a homodimer that contains 2 identical 2Fe-2S clusters." Biochemistry **33**(45): 13455-13463.
- Gaston, M. A., L. Zhang, et al. (2011). "The complete biosynthesis of the genetically encoded amino acid pyrrolysine from lysine." Nature **471**(7340): 647-U131.
- Giel, J. L., D. Rodionov, et al. (2006). "IscR-dependent gene expression links iron-sulphur cluster assembly to the control of O₂-regulated genes in *Escherichia coli*." Mol Microbiol **60**(4): 1058-1075.
- Glusker, J. P. (1980). "Citrate conformation and chelation: enzymic implications." Accounts Chem Res **13**(10): 345-352.
- Goodier, R. I. and B. M. M. Ahmer (2001). "SirA orthologs affect both motility and virulence." J Bacteriol **183**(7): 2249-2258.
- Gouet, P., E. Courcelle, et al. (1999). "ESPrpt: analysis of multiple sequence alignments in PostScript." Bioinformatics **15**(4): 305-308.
- Graham, D. E., H. M. Xu, et al. (2003). "Identification of the 7,8-didemethyl-8-hydroxy-5-deazariboflavin synthase required for coenzyme F₄₂₀ biosynthesis." Arch Microbiol **180**(6): 455-464.
- Han, S., R. S. Czernuszewicz, et al. (1989). "FE₂S₂ protein resonance Raman spectra revisited - Structural variations among adrenodoxin, ferredoxin, and red paramagnetic protein." J Am Chem Soc **111**(10): 3505-3511.
- Hanzelmann, P. and H. Schindelin (2006). "Binding of 5'-GTP to the C-terminal FeS cluster of the radical S-adenosylmethionine enzyme MoaA provides insights into its mechanism." Proc Natl Acad Sci USA **103**(18): 6829-6834.
- Hendrickson, E. L. and J. A. Leigh (2008). "Roles of coenzyme F₄₂₀-reducing hydrogenases and hydrogen- and F₄₂₀-dependent methylenetetrahydromethallopterin dehydrogenases in reduction of F₄₂₀ and production of hydrogen during methanogenesis." J Bacteriol **190**(14): 4818-4821.
- Hidese, R., H. Mihara, et al. (2011). "Bacterial cysteine desulfurases: versatile key players in biosynthetic pathways of sulfur-containing biofactors." Appl Microbiol Biot **91**(1): 47-61.
- Hiromoto, T., K. Ataka, et al. (2009). "The crystal structure of C176A mutated Fe -hydrogenase suggests an acyl- iron ligation in the active site iron complex." FEBS Lett **583**(3): 585-590.
- Holm, L. and P. Rosenstrom (2010). "Dali server: conservation mapping in 3D." Nucleic Acids Res **38**: W545-W549.

- Jacobson, M. R., K. E. Brigle, et al. (1989). "Physical and genetic map of the major nif gene cluster from *Azotobacter vinelandii*." *J Bacteriol* **171**(2): 1017-1027.
- Jacobson, M. R., V. L. Cash, et al. (1989). "Biochemical and genetic analysis of the nifUSVWZM cluster from *Azotobacter vinelandii*." *Mol Gen Genet* **219**(1-2): 49-57.
- Janin, J. and M. Delepierre (1999). *Biologie Structurale*, Hermann.
- Johnson, D. C., D. R. Dean, et al. (2005). Structure, function, and formation of biological iron-sulfur clusters. *Annu Rev Biochem* **74**: 247-281.
- Kabsch, W. (2010). "XDS." *Acta Crystallogr D* **66**: 125-132.
- Kaiser, J. T., T. Clausen, et al. (2000). "Crystal structure of a NifS-like protein from *Thermotoga maritima*: Implications for iron sulphur cluster assembly." *J Mol Biol* **297**(2): 451-464.
- Kantardjieff, K. A. and B. Rupp (2003). "Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals." *Protein Sci* **12**(9): 1865-1871.
- Kato, S., H. Mihara, et al. (2002). "Cys-328 of IscS and Cys-63 of IscU are the sites of disulfide bridge formation in a covalently bound IscS/IscU complex: Implications for the mechanism of iron-sulfur cluster assembly." *Proc Natl Acad Sci USA* **99**(9): 5948-5952.
- Kiley, P. J. and H. Beinert (2003). "The role of Fe-S proteins in sensing and regulation in bacteria." *Curr Opin Microbiol* **6**(2): 181-185.
- Kovacs, J. A., J. K. Bashkin, et al. (1985). "Persulfide-bridged iron molybdenum sulfur clusters of biological relevance - 2 synthetic routes and the structures of intermediate and product clusters." *J Am Chem Soc* **107**(6): 1784-1786.
- Krebs, C., J. N. Agar, et al. (2001). "IscA, an alternate scaffold for Fe-S cluster biosynthesis." *Biochemistry* **40**(46): 14069-14080.
- Kriek, M., F. Martins, et al. (2007). "Thiamine biosynthesis in *Escherichia coli*: Identification of the intermediate and by-product derived from tyrosine." *Angew Chem Int Edit* **46**(48): 9223-9226.
- Krissinel, E. and K. Henrick (2004). "Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions." *Acta Crystallogr D* **60**: 2256-2268.
- L'Haridon, S., V. Cilia, et al. (1998). "*Desulfurobacterium thermolithotrophum* gen. nov., sp. nov., a novel autotrophic, sulphur-reducing bacterium isolated from a deep-sea hydrothermal vent." *Int J Syst Bacteriol* **48**(3): 701-711.
- Langer, G., S. X. Cohen, et al. (2008). "Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7." *Nat Protoc* **3**(7): 1171-1179.
- Larkin, M. A., G. Blackshields, et al. (2007). "Clustal W and clustal X version 2.0." *Bioinformatics* **23**(21): 2947-2948.
- Lauble, H., M. C. Kennedy, et al. (1992). "Crystal structure of aconitase with isocitrate and nitroisocitrate bound." *Biochemistry* **31**(10): 2735-2748.
- Layer, G., J. Moser, et al. (2003). "Crystal structure of coproporphyrinogen III oxidase reveals cofactor geometry of Radical SAM enzymes." *EMBO J* **22**(23): 6214-6224.
- Lelivelt, M. J. and T. H. Kawula (1995). "HSC66, an HSP70 homolog in *Escherichia coli*, is induced by cold shock but not by heat-shock." *J Bacteriol* **177**(17): 4900-4907.
- Lepore, B. W., F. J. Ruzicka, et al. (2005). "The x-ray crystal structure of lysine-2,3-aminomutase from *Clostridium subterminale*." *Proc Natl Acad Sci USA* **102**(39): 13819-13824.
- Lill, R. (2009). "Function and biogenesis of iron-sulphur proteins." *Nature* **460**(7257): 831-838.
- Lima, C. D. (2002). "Analysis of the *E. coli* NifS CsdB protein at 2.0 angstrom reveals the structural basis for perselenide and persulfide intermediate formation." *J Mol Biol* **315**(5): 1199-1208.
- Liu, J. Y., N. Oganessian, et al. (2005). "Structural characterization of an iron-sulfur cluster assembly protein IscU in a zinc-bound form." *Proteins* **59**: 875-881.
- Lloyd, S. J., H. Lauble, et al. (1999). "The mechanism of aconitase: 1.8 angstrom resolution crystal structure of the S642A : citrate complex." *Protein Sci* **8**(12): 2655-2662.
- Lotierzo, M., E. Raux, et al. (2006). "Biotin synthase mechanism: Mutagenesis of the YNHNLD conserved motif." *Biochemistry* **45**(40): 12274-12281.
- Luscombe, N. M., S. E. Austin, et al. (2000). "An overview of the structures of protein-DNA complexes." *Genome Biol* **1**(1): reviews001.001 - reviews001.037.
- Marchler-Bauer, A., S. N. Lu, et al. (2011). "CDD: a Conserved Domain Database for the functional annotation of proteins." *Nucleic Acids Res* **39**: D225-D229.
- Marsh, E. N. G., D. P. Patterson, et al. (2010). "Adenosyl radical: reagent and catalyst in enzyme reactions." *Chembiochem* **11**(5): 604-621.
- Matthews, B. W. (1968). "Solvent content of protein crystals." *J Mol Biol* **33**(2): 491-&.

- McCoy, A. J., R. W. Grosse-Kunstleve, et al. (2007). "Phaser crystallographic software." *J Appl Crystallogr* **40**: 658-674.
- McGlynn, S. E., E. S. Boyd, et al. (2010). "Identification and characterization of a novel member of the radical AdoMet enzyme superfamily and implications for the biosynthesis of the Hmd hydrogenase active site cofactor." *J Bacteriol* **192**(2): 595-598.
- Meyer, J. (2008). "Iron-sulfur protein folds, iron-sulfur chemistry, and evolution." *J Biol Inorg Chem* **13**(2): 157-170.
- Mihara, H. and N. Esaki (2002). "Bacterial cysteine desulfurases: their function and mechanisms." *Appl Microbiol Biot* **60**(1-2): 12-23.
- Mihara, H., T. Kurihara, et al. (1997). "Cysteine sulfinatase desulfinase, a NIFS-like protein of *Escherichia coli* with selenocysteine lyase and cysteine desulfurase activities - Gene cloning, purification, and characterization of a novel pyridoxal enzyme." *J Biol Chem* **272**(36): 22417-22424.
- Mulder, D. W., E. M. Shepard, et al. (2011). "Insights into FeFe -Hydrogenase Structure, Mechanism, and Maturation." *Structure* **19**(8): 1038-1052.
- Murshudov, G. N., A. A. Vagin, et al. (1997). "Refinement of macromolecular structures by the maximum-likelihood method." *Acta Crystallogr D* **53**: 240-255.
- Nachin, L., L. Loiseau, et al. (2003). "SufC: an unorthodox cytoplasmic ABC/ATPase required for [Fe-S] biogenesis under oxidative stress." *EMBO J* **22**(3): 11.
- Nakamura, M., K. Saeki, et al. (1999). "Hyperproduction of recombinant ferredoxins in *Escherichia coli* by coexpression of the ORF1-ORF2-iscS-iscU-iscA-hscB-hscA-fdx-ORF3 gene cluster." *J Biochem* **126**(1): 10-18.
- Nicolet, Y., P. Amara, et al. (2009). "Unexpected electron transfer mechanism upon AdoMet cleavage in radical SAM proteins." *Proc Natl Acad Sci USA* **106**(35): 14867-14871.
- Nicolet, Y., A. L. de Lacey, et al. (2001). "Crystallographic and FTIR spectroscopic evidence of changes in Fe coordination upon reduction of the active site of the Fe-only hydrogenase from *Desulfovibrio desulfuricans*." *J Am Chem Soc* **123**(8): 1596-1601.
- Nicolet, Y. and C. L. Drennan (2004). "AdoMet radical proteins - from structure to evolution - alignment of divergent protein sequences reveals strong secondary structure element conservation." *Nucleic Acids Res* **32**(13): 4015-4025.
- Nicolet, Y., L. Martin, et al. (2010). "A glycol free radical as the precursor in the synthesis of carbon monoxide and cyanide by the FeFe -hydrogenase maturase HydG." *FEBS Lett* **584**(19): 4197-4202.
- Nicolet, Y., J. K. Rubach, et al. (2008). "X-ray structure of the FeFe -hydrogenase maturase HydE from *Thermotoga maritima*." *J Biol Chem* **283**(27): 18861-18872.
- Numata, T., S. Fukai, et al. (2006). "Structural basis for sulfur relay to RNA mediated by heterohexameric TusBCD complex." *Structure* **14**(2): 357-366.
- Nuth, M., T. Yoon, et al. (2002). "Iron-sulfur cluster biosynthesis: Characterization of iron nucleation sites for assembly of the 2Fe-2S (2+) cluster core in IscU proteins." *J Am Chem Soc* **124**(30): 8774-8775.
- Ollagnier-de Choudens, S., Y. Sanakis, et al. (2002). "Reductive cleavage of S-adenosylmethionine by biotin synthase from *Escherichia coli*." *J Biol Chem* **277**(16): 13449-13454.
- Ollagnier-De-Choudens, S., D. Lascoux, et al. (2003). "Mechanistic studies of the SufS-SufE cysteine desulfurase: evidence for sulfur transfer from SufS to SufE." *FEBS Lett* **555**(2): 263-267.
- Olson, B. J. S. C. and J. Markwell (2001). Assays for determination of protein concentration. *Current Protocols in Pharmacology*, John Wiley & Sons, Inc.
- Outten, F. W., O. Djaman, et al. (2004). "A suf operon requirement for Fe-S cluster assembly during iron starvation in *Escherichia coli*." *Mol Microbiol* **52**(3): 861-872.
- Parey, K., E. Warkentin, et al. (2010). "Reaction cycle of the dissimilatory sulfite reductase from *Archaeoglobus fulgidus*." *Biochemistry* **49**(41): 8912-8921.
- Pilet, E., Y. Nicolet, et al. (2009). "The role of the maturase HydG in FeFe -hydrogenase active site synthesis and assembly." *FEBS Lett* **583**(3): 506-511.
- Posewitz, M. C., P. W. King, et al. (2004). "Discovery of two novel radical S-adenosylmethionine proteins required for the assembly of an active Fe hydrogenase." *J Biol Chem* **279**(24): 25711-25720.
- Pott, A. S. and C. Dahl (1998). "Sirohaem sulfite reductase and other proteins encoded by genes at the dsr locus of *Chromatium vinosum* are involved in the oxidation of intracellular sulfur." *Microbiol-UK* **144**: 1881-1894.
- Prischi, F., P. V. Konarev, et al. (2010). "Structural bases for the interaction of frataxin with the central components of iron-sulphur cluster assembly." *Nat Commun* **1**.
- Prischi, F., C. Pastore, et al. (2010). "Of the vulnerability of orphan complex proteins: The case study of the *E. coli* IscU and IscS proteins." *Protein Express Purif* **73**(2): 161-166.

- Py, B. and F. Barras (2010). "Building Fe-S proteins: bacterial strategies." *Nat Rev Micro* **8**(6): 436-446.
- Quitterer, F., A. List, et al. (2011). "Crystal Structure of Methylornithine Synthase (PylB): Insights into the Pyrrolysine Biosynthesis." *Angew Chem Int Edit*: n/a-n/a.
- Ramachandran, G. N., C. Ramakrishnan, et al. (1963). "Stereochemistry of polypeptide chain configurations." *J Mol Biol* **7**(1): 95-&.
- Raulfs, E. C., I. P. O'Carroll, et al. (2008). "In vivo iron-sulfur cluster formation." *Proc Natl Acad Sci USA* **105**(25): 8591-8596.
- Rubach, J. K., X. Brazzolotto, et al. (2005). "Biochemical characterization of the HydE and HydG iron-only hydrogenase maturation enzymes from *Thermatoga maritima*." *FEBS Lett* **579**(22): 5055-5060.
- Rubio, L. and P. W. Ludden (2005). "Maturation of nitrogenase: a biochemical puzzle." *J Bacteriol* **187**(2): 10.
- Rubio, L. M. and P. W. Ludden (2008). Biosynthesis of the Iron-Molybdenum Cofactor of Nitrogenase. *Annu Rev Microbiol.* **62**: 93-111.
- Schleucher, J., C. Griesinger, et al. (1994). "H-2-forming N-5,N-10-methylenetetrahydromethanopterin dehydrogenase from *Methanobacterium thermoautotrophicum* catalyzes a stereoselective hydride transfer as determined by 2-dimensional NMR spectroscopy." *Biochemistry* **33**(13): 3986-3993.
- Schrödinger (2011). "Maestro v 9.2, Prime v 3.0, QSite v 5.6, Schrödinger, LLC, New York, NY."
- Schwartz, C. J., O. Djaman, et al. (2000). "The cysteine desulfurase, IscS, has a major role in in vivo Fe-S cluster formation in *Escherichia coli*." *Proc Natl Acad Sci USA* **97**(16): 9009-9014.
- Schwartz, C. J., J. L. Giel, et al. (2001). "IscR, an Fe-S cluster-containing transcription factor, represses expression of *Escherichia coli* genes encoding Fe-S cluster assembly proteins." *Proc Natl Acad Sci USA* **98**(26): 14895-14900.
- Shepard, E. M., B. R. Duffus, et al. (2010). "FeFe -Hydrogenase maturation: HydG-catalyzed synthesis of carbon monoxide." *J Am Chem Soc* **132**(27): 9247-9249.
- Shi, R., A. Proteau, et al. (2010). "Structural basis for Fe-S Cluster assembly and tRNA thiolation mediated by IscS protein-protein interactions." *Plos Biol* **8**(4).
- Shima, S. and U. Ermler (2011). "Structure and function of Fe -Hydrogenase and its Iron-Guanylylpyridinol (FeGP) Cofactor." *Eur J Inorg Chem*(7): 963-972.
- Shimomura, Y., H. Kamikubo, et al. (2007). "Characterization and crystallization of an IscU-type scaffold protein with bound 2Fe-2S cluster from the hyperthermophile, *Aquifex aeolicus*." *J Biochem* **142**(5): 577-586.
- Shimomura, Y., K. Wada, et al. (2008). "The asymmetric trimeric architecture of 2Fe-2S IscU: Implications for its scaffolding during iron-sulfur cluster biosynthesis." *J Mol Biol* **383**(1): 133-143.
- Smith, A. D., J. N. Agar, et al. (2001). "Sulfur transfer from IscS to IscU: The first step in iron-sulfur cluster biosynthesis." *J Am Chem Soc* **123**: 11103-11104.
- Smith, A. D., G. N. L. Jameson, et al. (2005). "NifS-mediated assembly of 4Fe-4S clusters in the N- and C-terminal domains of the NifU scaffold protein." *Biochemistry* **44**(39): 12955-12969.
- Sofia, H. J., G. Chen, et al. (2001). "Radical SAM, a novel protein superfamily linking unresolved steps in familiar biosynthetic pathways with radical mechanisms: functional characterization using new analysis and information visualization methods." *Nucleic Acids Res* **29**(5): 1097-1106.
- Stetter, K. O., G. Lauerer, et al. (1987). "Isolation of extremely thermophilic sulfate reducers - evidence for a novel branch of archaebacteria." *Science* **236**(4803): 822-824.
- Suzuki, Y., A. Noma, et al. (2007). "Crystal structure of the radical SAM enzyme catalyzing tricyclic modified base formation in tRNA." *J Mol Biol* **372**(5): 1204-1214.
- Takahashi, Y. and M. Nakamura (1999). "Functional assignment of the ORF2-iscS-iscU-iscA-hscB-hscA-fdx-ORF3 gene cluster involved in the assembly of Fe-S clusters in *Escherichia coli*." *J Biochem* **126**(5): 917-926.
- Takahashi, Y. and U. Tokumoto (2002). "A third bacterial system for the assembly of iron-sulfur clusters with homologs in archaea and plastids." *J Biol Chem* **277**(32): 28380-28383.
- Taylor, G. L. (2010). "Experimental phasing and radiation damage - Introduction to phasing." *Acta Crystallogr D* **66**: 325-338.
- Thauer, R. K., A.-K. Kaster, et al. (2010). "Hydrogenases from methanogenic Archaea, Nickel, a novel cofactor, and H(2) Storage." *Annu Rev Biochem* **79**: 507-536.
- Tirupati, B., J. L. Vey, et al. (2004). "Kinetic and structural characterization of Slr0077/SufS, the essential cysteine desulfurase from *Synechocystis* sp PCC 6803." *Biochemistry* **43**(38): 12210-12219.
- Tokumoto, U. and Y. Takahashi (2001). "Genetic analysis of the isc operon in *Escherichia coli* involved in the biogenesis of cellular iron-sulfur protein." *J Biochem* **130**(1): 63-71.
- Tron, C., M. V. Cherrier, et al. (2011). "Further characterization of the FeFe -Hydrogenase maturase HydG." *Eur J Inorg Chem*(7): 1121-1127.

- Trotter, V., D. Vinella, et al. (2009). "The CsdA cysteine desulphurase promotes Fe/S biogenesis by recruiting Suf components and participates to a new sulphur transfer pathway by recruiting CsdL (ex-YgdL), a ubiquitin-modifying-like protein." *Mol Microbiol* **74**(6): 1527-1542.
- Unciuleac, M. C., K. Chandramouli, et al. (2007). "*In vitro* activation of apo-aconitase using a 4Fe-4S cluster-loaded form of the IscU Fe-S cluster scaffolding protein." *Biochemistry* **46**(23): 6812-6821.
- Urbina, H. D., J. J. Silberg, et al. (2001). "Transfer of sulfur from IscS to IscU during Fe/S cluster assembly." *J Biol Chem* **276**: 44521-44526.
- Vernede, X. and J. C. Fontecilla-Camps (1999). "A method to stabilize reduced and or gas-treated protein crystals by flash-cooling under a controlled atmosphere." *J Appl Crystallogr* **32**: 505-509.
- Vey, J. L. and C. L. Drennan (2011). "Structural insights into radical generation by the radical SAM superfamily." *Chem Rev* **111**(4): 2487-2506.
- Vey, J. L., J. Yang, et al. (2008). "Structural basis for glycyl radical formation by pyruvate formate-lyase activating enzyme." *Proc Natl Acad Sci USA* **105**(42): 16137-16141.
- Vignais, P. M. and B. Billoud (2007). "Occurrence, classification, and biological function of hydrogenases: An overview." *Chem Rev* **107**(10): 4206-4272.
- Wada, H. and E. E. Snell (1962). "Enzymatic transamination of pyridoxamine. I. With oxaloacetate and alpha-ketoglutarate." *J Biol Chem* **237**: 127-132.
- Walters, A. D. and J. P. J. Chong (2009). "*Methanococcus maripaludis*: an archaeon with multiple functional MCM proteins?" *Biochem Soc T* **37**: 1-6.
- Wang, B. C. (1985). "Resolution of phase ambiguity in macromolecular crystallography." *Method Enzymol* **115**: 90-112.
- Wood, H. E., K. M. Devine, et al. (1990). "Characterization of a repressor gene (xre) and a temperature-sensitive allele from the *Bacillus subtilis* prophage, PbsX." *Gene* **96**(1): 83-88.
- Wu, G., S. S. Mansy, et al. (2002). "Characterization of an iron-sulfur cluster assembly protein (ISU1) from *Schizosaccharomyces pombe*." *Biochemistry* **41**(15): 5024-5032.
- Yuvaniyama, P., J. N. Agar, et al. (2000). "NifS-directed assembly of a transient 2Fe-2S cluster within the NifU protein." *Proc Natl Acad Sci USA* **97**(2): 599-604.
- Zhang, Q., D. Chen, et al. (2011). "Characterization of NocL involved in thiopeptide Nocathiacin I biosynthesis: a 4Fe-4S cluster and the catalysis of a radical S-Adenosylmethionine enzyme." *J Biol Chem* **286**(24): 21287-21294.
- Zhang, Q., Y. Li, et al. (2011). "Radical-mediated enzymatic carbon chain fragmentation-recombination." *Nat Chem Biol* **7**(3): 154-160.
- Zhang, W. J., A. Urban, et al. (2010). "IscS functions as a primary sulfur-donating enzyme by interacting specifically with MoeB and MoeD in the biosynthesis of molybdopterin in *Escherichia coli*." *J Biol Chem* **285**(4): 2302-2308.
- Zhang, Y., X. Zhu, et al. (2010). "Diphthamide biosynthesis requires an organic radical generated by an iron-sulphur enzyme." *Nature* **465**(7300): 891-U894.
- Zheng, L., R. H. White, et al. (1994). "Mechanism for the desulfurization of L-cysteine catalyzed by the nifS gene product." *Biochemistry* **33**: 4714-4720.
- Zheng, L. M., V. L. Cash, et al. (1998). "Assembly of iron-sulfur clusters - Identification of an iscSUA-hscBA-fdx gene cluster from *Azotobacter vinelandii*." *J Biol Chem* **273**(21): 13264-13272.
- Zheng, L. M. and D. R. Dean (1994). "Catalytic formation of a nitrogenase iron-sulfur cluster." *J Biol Chem* **269**(29): 18723-18726.
- Zheng, L. M., R. H. White, et al. (1993). "Cysteine desulfurase activity indicates a role for Nifs in metallocluster biosynthesis." *Proc Natl Acad Sci USA* **90**: 2754-2758.